Statistical Analysis of Moose Habitat Behaviors Using Bayesian Hierarchical Model with Spatially Varying Coefficients

A PROJECT SUBMITTED TO THE FACULTY OF UNIVERSITY OF MINNESOTA BY

Matěj Kroc

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

Xuan Li

June, 2017

Acknowledgments

Firstly, I would like to thank my advisors Dr. Xuan Li for all of his time, guidance and expert support while completing this project. I would also like to thank Dr. Richard Green and Dr. Ron Moen for serving on my committee. Secondly, my thanks should also be extended to the faculty and staff of the department of Mathematics and Statistics at University of Minnesota Duluth for everything they have provided me with and helped me accomplish over the last two years. Lastly, I would like to give special thanks to my family and closest friends for their support, encouragement and help as I studied abroad.

Abstract

In the past few years the Bayesian spatial models are becoming increasingly important due to the higher ability to collect larger amounts of spatial data in various domains. The goal of this project is to utilize a Bayesian hierarchical model with spatially varying coefficients to obtain better insights into moose habitat behavior in Northern Minnesota with emphasis upon the significance of spatial coefficients. The spatial Bayesian model was fitted on the sampled moose data and compared with two non-spatial models. Model framework, parameter specification and Bayesian inference were discussed. Results showed that the addition of spatially varying coefficients improved model performance.

Contents

List of Figures

List of Tables

1 Introduction

In the past few years interest in statistical modeling has rapidly increased for scientists in many different fields. With new technologies and the ability to collect larger amounts of data they sought a tool which would help them to get a better understanding, and eventually, prediction of behavior of subjects in their range of study. For biologists and ecologists habitat data is necessary to develop effective conservation and management strategies, and help determine what is behind the change in the population of different species.

Our research is focused on the moose habitat behavior statistics. Moose, *Alces alces*, are the largest of all deer species. Male moose are recognizable by their huge antlers, which can spread up to 6 feet wide. Because of their tall body, they prefer to browse higher shrubs and their typical habitat is a dense mixed boreal forest in North America, including the northern United States, Canada, Alaska, and in Scandinavia and Russia. Despite their large bodies, moose are good swimmers and are often seen in lakes and rivers feeding on aquatic plants both at and below the surface [15]. One of the reasons why moose habitat behavior is the subject of study by many biologists is recent changes in population in North America. Since the 1990's, the moose population in northern Minnesota has decline significantly. Based on a moose population survey from 2017 [16], the population in northeastern Minnesota has dropped from about $8,000$ moose to a stable population of just under $4,000$ moose over the last 4 years. Meanwhile, the northwestern Minnesotan population practically disappeared after declining from $4,000$ to fewer than 100. The reason behind this steep drop is unknown. Many scientists believe that it could be caused by climate change. Shorter winters and longer falls give more time for parasites, especially winter ticks, to find a host. For purposes of research, moose wore GPS collars, which allow biologists to track their location and collect essential data for future work. In some cases, moose received a tiny transmitter which monitored their heart rate and temperature and notified biologists when the moose died [17].

This work intends to utilize the Bayesian hierarchic model with spatially varying coefficients to obtain better insights into moose habitat behavior in Northern Minnesota.

Our goal is to fit, and eventually predict, moose behavior connected with seeking fresh water based on the data collected by GPS collars. Use of Bayesian hierarchical models has received much attention in the last few years. This is mainly due to the increasing availability of spatial data in various domains. In the classic regression model we do not account for spatial dependence of residuals. A non-spatial model can be appropriate if all spatially structured variations in the outcome are accounted for by the covariates used for model fitting. However, this is often an unrealistic assumption when data are spatially indexed. There is a reason to expect a similar outcome for neighboring locations. In this paper we will also employ a Generalized linear model and a Bayesian Generalized linear regression model as two non-spatial models to serve as a comparison with Bayesian spatial regression model to evaluate the necessity of a spatial component. For the spatial modeling we will use a Univariate Bayesian Generalized linear spatial regression model.

The remainder of the paper evolves as follows. The motivating data and preprocessing steps are detailed in Section 2. Brief theory behind the Bayes' modeling is described in section 3. The proposed modeling framework and model assessment are described in Section 4, followed by analysis results presented in Section 5. Finally, some discussion along with an indication of future work is provided in Section 6.

2 Data Description

This section is devoted to data introduction and some summary statistics. We describe how data were collected and present the individual variables. We also more closely focus on the variables we are going to use to fit our model and their distribution.

2.1 GPS Data

Field data were collected using GPS collars. Each moose wore one collar and data were recorded in 20 minute intervals. Every observation consists of a unique ID, ID of the moose, date, time, temperature of the collar in degrees Celsius, vegetation description, UTMx and UTMy coordinates. The last data field is Hydro, which indicates if a body of water is present nearby. The data set contains $213,887$ data points for 10 moose. Since the location ID is unique, it ranges from 1 to $213,887$. Not all of the moose have the same amount of observations. Number of observations for each moose based on their ID is presented in table 1. Collar locations in this dataset are dated from $4/15$ to $10/31$ in years 2011 and 2012. However, some moose were tracked just one year. Each period includes exactly 200 days. Responses were collected in 20 minute intervals, thus covariate time consists of 72 different values. Since we have 72 20-minute periods for each day and total of 200 or 400 days were measured, the frequency for each moose ID should be $14,400$ or $28,800$. None of the frequencies in table 1 reach this number; therefore, some of the observations for each moose are missing. This is caused by moose losing their collars, a malfunction of the collar, as a missed GPS locations. UTM (Universal Transverse Mercator) coordinates are used to record locations. A pair of variables UTMx (Northing) and UTMy (Easting) gives the location on the landscape. In the UTM coordinate system, the Earth's surface is divided into 16 zones. Since our data were collected in the northern Minnesotan region, the zone number is 15. Also for each of the recorded locations, description of vegetation ("Open water", "Mixedwood forest", etc.) is included along with its numerical class code obtained by satellite screening.

The variable we would like to use in our model as a response is called Hydro. In the original dataset Hydro has two outcomes: 1) an empty cell (N/A) or 2) string "PRESENT" is collected as a response. This variable records the locations when a moose (the GPS collar) was with in close distance to water or in the water. Therefore, observations marked as Hydro "PRESENT" are locations where a moose was most likely drinking water, swimming, or waiting in the water and feeding on aquatic plants. According to the variable Hydro, 1.95% of all the locations were recognized as "PRESENT".

Moose_ID	Frequency	Percent recorded
31166	28453	98.80
31168	14393	99.95
31169	28453	98.80
31170	14391	99.94
31172	14062	97.65
31174	28458	98.81
31178	28450	98.79
31179	14390	99.93
31182	14390	99.93
31184	28447	98.77

Table 1: Frequency of observations for each moose ID. Percent recorded shows amount of locations successfully recorded from total of $28,800$ or $14,400$.

2.2 Ely Weather Station Data

For each observation, the data set includes measurements from Ely (MN) weather station. This includes temperature, relative humidity, dew point and time of sunrise and sunset. Dew point is the temperature at which air must be cooled to become saturated with water vapor. Table 2 provides a summary for collar and Ely weather station temperatures. The collar temperature varies from -5°C to 54°C and it has no missing values. Ely temperature has an average of 13.08°C with minimum and maximum values of -12°C and 33°C, respectively. None of the variables have any missing values.

Variable	Min	Mean	Max	Std. deviation
Collar Temp	-5	17.64	54	7.46
Ely Temp	-12	13.08	33	8.19

Table 2: Summary of collar and Ely weather station temperatures.

2.3 Hydro PRESENT

We are interested in observations where covariate HYDRO has a value “PRESENT”, indicating the moose were in or near water. By looking at figure 1, where we illustrate the portion of hydro "PRESENT" observations according to months, we can see that there is a significant difference among percentages for April and June and the other five months. Low percentage for the month of April could be explained by the lowest average temperature, but especially by the fact that data were collected only for 4/15 – 4/30, thus there is generally fewer observations for April in our dataset. Regarding June, mean collar and Ely temperature were 20.64°C and 16.20°C, respectively, which is unexpectedly not the highest average we observed if we expect the highest occurrence of Hydro “PRESENT” connected with the higher temperatures. The highest mean temperature is for July, 23.70°C. However, these temperatures are not much different and mean temperatures tend to be higher if we look specifically at observations where the location is in a water feature. Temperature rises about one degree for both months. Thus, a better explanation of higher activity around water for June could possibly be feeding since it is a aquatic plants season, also it is a typical calving season, along with May, for moose.

The last covariate from our dataset is vegetation class description. This variable shows the vegetation type for each location and it is represented by a verbal description such as: “Coniferous forest”, “Bog”, “Open water”, etc. For statistical modeling purposes these class names are connected with a categorical variable “class code” which represents each class as an integer. Based on this covariate, the most observations were located in “Mixedwood forest” and “Regeneration/young forest”. The frequencies are $98,951$ (46.26%) and $39,005$ (18.24%), respectively.

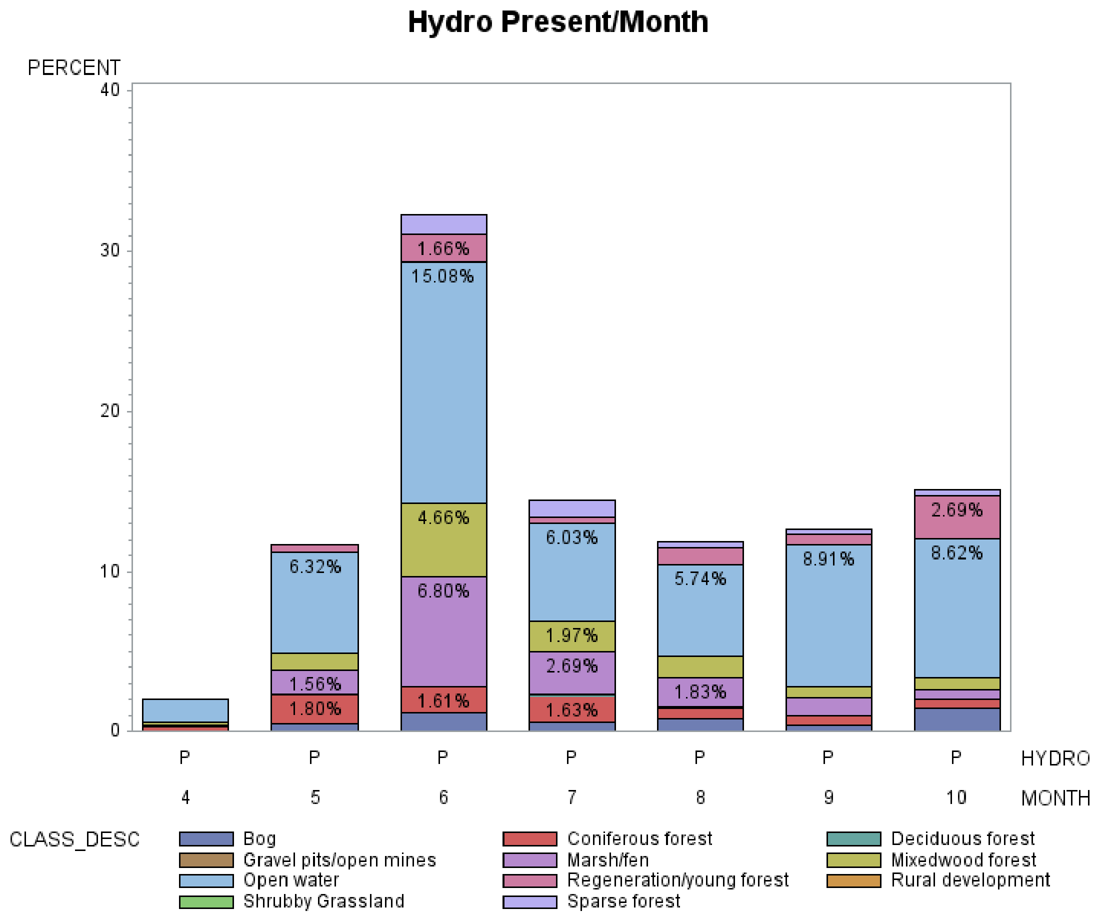

Figure 1: Presence of water according to months sub-grouped by vegetation description.

Another aspect of figure 1 is that it shows the vegetation description for each location where hydro was measured as present. We can see that the highest percentage occurs at locations with open water. This is definitely one of the expected outcomes since open water such as lakes, ponds and water basins are the biggest source of fresh water for moose. Since the variable Hydro represents the situation when the collar was in the close proximity to water. Thus in the whole dataset there are 4,986 locations with the class code 3 ("Open water") but only 2,170 are also marked as hydro "PRESENT". Hence only in 43.53% of cases moose were near the larger open water. The second and third largest represented classes are "Marsh/fen" and "Mixedwood forest" with 614 (13.26%) and 441 (0.45%), respectively, for locations marked as hydro "Present". As we mentioned before, low total percentage for mixed wood forest is due to the fact it is the most common vegetation type in our dataset.

In figure 2, we see the points collected by the GPS collar for moose with ID number 31168. Range of spread of the locations for all moose is $\sim$1 m to $7,000$ m. This information is further used as an effective spatial range for determining the support of a spatial decay parameter.

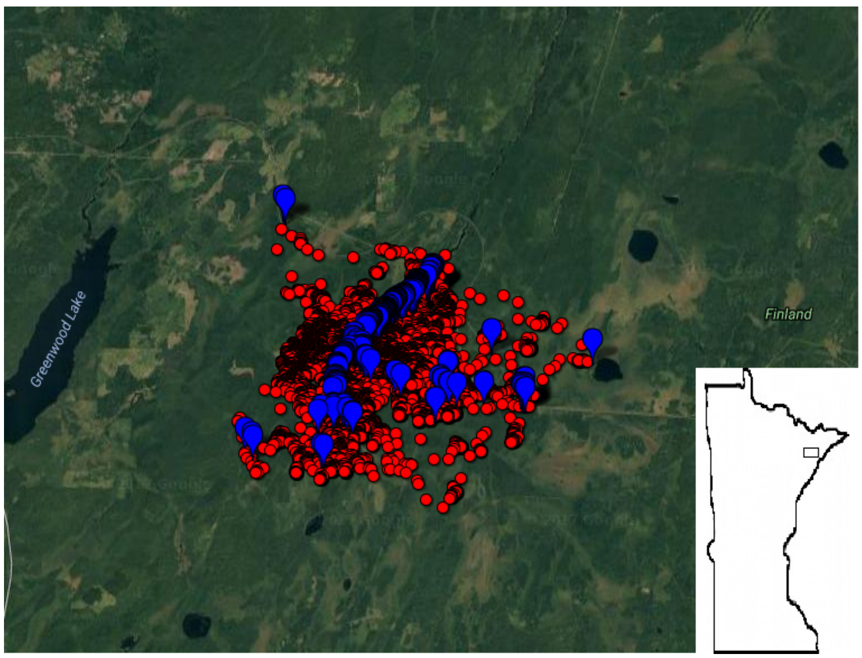

Figure 2: Recorded location of moose with ID number 31168. Locations which were recognized as they are in the water feature are marked by blue symbols.

2.4 New Variables

In order to improve the fit and prediction performance of our model we decided to create some new variables. We do not change or eliminate any observation. First of all, we want to model the variable Hydro, which is of type string; therefore, we simply introduce a new variable Hydro_pom where response N/A (empty cell) and "PRESENT" have outcomes 0 and 1, respectively.

Secondly, we would like to involve a variable which would help us to observe if there is a difference in behavior between day and night. Thus, we use the variable Time for each location and time of sunrise and sunset at Ely weather station to introduce the new variable, Time_NEW, where outcome equals 0 if location time is recorded before sunrise or after sunset and equals 1 otherwise.

$$Hydro\_pom = \begin{cases} 1 & if \quad Hydro = PRESENT \\ 0 & otherwise \end{cases}$$

$$Time\_NEW = \begin{cases} 1 & if \quad sunrise \leq Time \leq sunset \\ 0 & otherwise \end{cases}$$

Thirdly, we would like to reduce the number of different classes of vegetation to reduce complexity. Therefore we used the following variant of class code distribution (on the right-hand side we have a new purposed class code number and on the left-hand side merged vegetation classes with their description and old class code in parenthesis):

- Class code 1: Open water (3),
- Class code 2: Marsh/Fen (6) + Bog (7),
- Class code 3: Mixedwood forest (5) + Coniferous (9) + Deciduous (2),
- Class code 4: Regeneration(14)+Sparse forest (99),
- Class code 5: Gravel pits (12)+Rural development (10)+Shrubby grassland (11).

After the regrouping we noticed that there are very few observations for group number 5

(255 locations) compared to the other groups; therefore, we decided to treat these sites as outliers and remove them from the data set. None of these observations were marked as Hydro "PRESENT". We handle this variable as a categorical variable with group number 2 as a reference (contrast) variable.

Last but not least, we would like to observe how different times throughout April to October effect our outcome variable Hydro. Thus we created categorical variable `Date_NEW`, which splits the observation period between April and October into three groups described below. The group number 2 includes the time of the year with the highest temperatures and we use this group as a reference variable. Group assignment is the same for both years 2011 and 2012.

$$Date\_NEW = \begin{cases} 1 \quad for \quad 04/15 - 05/31 \\ 2 \quad for \quad 06/01 - 07/31 \\ 3 \quad for \quad 08/01 - 10/31 \end{cases}$$

As we will be able to see later, these newly purposed variables helped us to improve the model fit performance for a Generalized linear model (GLM) based on the Akaike information criterion (AIC) statistics and Bayesian models based on the Deviance information criterion (DIC).

2.5 Covariate Selection

Our dataset includes several different variables as mentioned in previous descriptions. We could use all of them as predictors for our models, however application of some of them would be useless in a sense that these variables would not have much impact on model fitting or they would be eliminated by backward selection. In this section we would like to provide the reasoning for our final variable selection.

First, let us remove variables which are not meant to be used in the final fitting or were used to create new variables in section 2.4. Thus, we remove: `Moose_ID, sunrise_time, sunset_time` and `Hydro`. We will also remove other variables we will not use as covariates are coordinates `UTMx` and `UTMy` since these will later be provided to the spatial model as locations, therefore they are not treated as covariates of the model.

Next, in the last section 2.4 we created new variables; thus, it would not be appropriate to use two covariates for time or class code. Based on the tests on all observations for GLM and random sample with n observations for Bayesian models, we decided to adopt as covariates `Time_NEW` and `CLASS_CODE_NEW` since they showed better model fitting.

Last but not least, we have variables for collar temperature, Ely weather station temperature, relative humidity and dew point temperature. Ely temperature represents in our data set the daily temperatures and as we may see in Table 3, the other three variables correlate with this temperature. Therefore, we decided not to employ either of `Ely_RH` nor `Ely_DWPT`. We decided to go this way not just because of the high correlation, but also because of uncertainty in the usefulness of these variables in predicting the right outcome. Lastly, there is high correlation between collar and Ely temperatures. The collar temperatures are measured spatially by the GPS collar in the exact location, therefore we chose the variable `Collar_Temp` as a covariate for our model, since it includes the spatial component.

	Collar_Temp	**Ely_Temp**	**Ely_RH**	**Ely_DWPT**
Collar_Temp	1.000	0.898	−0.501	0.599
Ely_Temp	0.898	1.000	−0.481	0.733
Ely_RH	−0.501	−0.481	1.000	0.227
Ely_DWPT	0.599	0.733	0.227	1.000

Table 3: Correlation among the variables associated with temperature.

3 Bayesian inference

In this chapter we provide a brief review of Bayesian hierarchic modeling and the basic principles of Bayesian inference.

3.1 Bayes' Theorem

The basic problem in scientific/ecological research is to understand processes that can not be observed based on quantities that we can observe. We represent unobserved processes as models made up of parameters and latent states, which we notate here as $\boldsymbol{\theta}$. In this approach, in addition to specifying the distributional model $f(\boldsymbol{y}|\boldsymbol{\theta})$ for the observed data $\boldsymbol{y} = (y_1, ..., y_n)$ given a vector of unknown parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)$, we assume that $\boldsymbol{\theta}$ is a random variable governed by the probability distribution $\pi(\boldsymbol{\theta})$. We call this distribution a *prior* distribution [2]. The probability function $f(\boldsymbol{y}|\boldsymbol{\theta})$ is called *likelihood* and is defined as

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} f(y_i, \boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)$ and $f(\boldsymbol{y}, \boldsymbol{\theta})$ is a probability density function. We wish to discover the probability distribution of the unobserved $\boldsymbol{\theta}$ conditional on the observed data $\boldsymbol{y}$, that is $p(\boldsymbol{\theta}|\boldsymbol{y})$, the *posterior* distribution. Using the basic rules of conditional probability for two random variables, we have

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \frac{f(\boldsymbol{y}, \boldsymbol{\theta})}{\pi(\boldsymbol{\theta})}, \tag{2}$$

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{f(\boldsymbol{y}, \boldsymbol{\theta})}{f_y}. \tag{3}$$

By solving equation 2 for $f(\boldsymbol{y}, \boldsymbol{\theta})$ and substituting the right-hand side in equation 3 we get

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{f_y}, \tag{4}$$

where f_y is a marginal function of $f(\boldsymbol{y}, \boldsymbol{\theta})$ integrated over $\boldsymbol{\theta}$. Hence

$$f_y = \int_{-\infty}^{\infty} f(\boldsymbol{y}, \boldsymbol{\theta})d\boldsymbol{\theta} = \int_{-\infty}^{\infty} f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}, \tag{5}$$

thus by substituting in 4 we obtain Bayes' Theorem for parameters that are continuous as

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{-\infty}^{\infty} f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{6}$$

Therefore we obtained the *posterior* distribution as a product of the *likelihood* and *prior* over the marginal distribution of $f(\boldsymbol{y}, \boldsymbol{\theta})$. Dividing the joint distribution by $\int_{-\infty}^{\infty} f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ assures that the posterior distribution integrates to 1. Because of this, f_y is often called a *normalizing* constant [1], [2]. Before the data are collected, f_y is an unknown random variable, however, after the data are collected, f_y is known, fixed and this distribution can be specified as a proportionality

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto f(\boldsymbol{y}, \boldsymbol{\theta}) \propto f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}). \tag{7}$$

3.2 Hierarchical Bayesian Models

In the previous subsection we introduced Bayes' Theorem and showed a simple Bayesian model since it represents the joint distribution of random variables whose probability distribution is estimated from observations and prior knowledge, as the product of likelihood and prior distributions. In a hierarchical model the prior distribution for the parameter also has parameters controlling its form. The distributions specified for these parameters are known as hyperprior distributions, and the parameters are known as hyperparameters. Thus, the prior distribution is written as $\pi(\boldsymbol{\theta}|\boldsymbol{\lambda})$, where $\boldsymbol{\lambda}$ is a vector of hyperparameters

[1], [6]. Therefore the proportionality 7 would be replaced by

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto f(\boldsymbol{y}, \boldsymbol{\theta}) \propto f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{\lambda})h(\boldsymbol{\lambda}), \tag{8}$$

where $h(\boldsymbol{\lambda})$ are hyperprior distributions.

In other words, we can separate the model into three stages. The first stage is the data model, where we have the "true" ecological underlying process $\boldsymbol{\theta}$ which is not observable. This state relates on the observable data $\boldsymbol{y}$, using a model with a vector of parameters $\boldsymbol{\theta}$ and $\boldsymbol{\theta_0}$. Stage two, the process model, predicts the behavior of the true underlying process based on a vector of parameters $\boldsymbol{\lambda}$, the hyperpriors. Lastly, stage three, the parameter model, includes the data and true process parameters $\boldsymbol{\theta_0}$ and $\boldsymbol{\lambda}$ [7]. Thus:

- Stage 1: Data model [data| "true" process, data parameters] $\rightarrow [\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\theta_0}]$
- Stage 2: Process model [process| process parameters] $\rightarrow [\boldsymbol{\theta}|\boldsymbol{\lambda}]$
- Stage 3: Parameter model [data and process parameters] $\rightarrow [\boldsymbol{\theta_0}], [\boldsymbol{\lambda}]$

3.3 Bayesian Computation

In previous sections we showed how to use the Bayes' theorem to obtain a Bayesian simple and hierarchical model. We also stated that our goal is to discover a probability distribution of the unknown underlying process $\boldsymbol{\theta}$ conditional on the observed data $\boldsymbol{y}$. In this subsection, we would like to briefly explain how to obtain the joint posterior distribution.

Recall equation 6, the Bayes' Theorem to obtain $p(\boldsymbol{\theta}|\boldsymbol{y})$ includes the integral of the joint distribution taken over all the parameters. For the most simple problem we can use a straightforward approach, however, in most of the models with more parameters it is impossible to integrate the joint distribution analytically. Therefore, we need to find a way to avoid formal integration. The most widely used computing tools in Bayesian practice today are Markov Chain Monte Carlo (MCMC) algorithms. MCMC allows us to find the marginal distribution of each of the unknown variables. MCMC methods are a class of algorithms for sampling from a probability distribution based on constructing a Markov chain that has the desired distribution as its equilibrium distribution, in our case $p(\boldsymbol{\theta}|\boldsymbol{y})$. The state of the chain after a number of steps is then used as a sample of the desired distribution. The quality of the sample improves as a function of the number of steps [1], [6]. The two most common algorithms are the Gibbs Sampler and the Metropolis-Hastings algorithm. These two MCMC algorithms are implemented in R packages *MCMCpack* [8] and *spBayes* [9], which we are going to use later to obtain posterior samples for our proposed Bayesian models.

R package *spBayes* [9] offers two options of implementation of MCMC algorithms. There is the traditional random-walk Metropolis-Hastings algorithm to update the spatial effects. Alternately, an adaptive MCMC Metropolis-within-Gibbs algorithm, proposed by Roberts and Rosenthal [11], is available for a more automated function call. For a traditional option we must specify Metropolis proposal variances, i.e., *tuning* values and we then monitor acceptance rates for the parameters and possibly change the *tuning* values to obtain the desired acceptance rate. For the adaptive method we specify a number and a size of batches and a desired overall acceptance rate. For simplicity, we will always implement the MCMC using the adaptive method; therefore, we do not have to specify the *tuning* values for our models. We can observe the proposed tuning in the algorithm report for each batch.

4 Models

The intention of this section is to introduce the regression model with spatially varying coefficients which will be used to model the outcome variable y.

4.1 Model Framework

Since we would like to model the variable Hydro as a dependent outcome variable, for these purposes we will use a non-Gaussian model, since, recall now from subsection 2.1 where we introduced the collected GPS data, our outcome variable is not continuous, but binary. In the non-Gaussian model we replace the Gaussian likelihood function with the assumption that $E[y(\boldsymbol{s})]$ is linear on a transformed scale, thus

$$g(E[y(\boldsymbol{s})]) = \boldsymbol{x(s)^T\beta} + \omega(\boldsymbol{s}), \tag{9}$$

where $g(\cdot)$ is a suitable link function, which provides the relationship between the linear predictor and the mean of the distribution function [10]. We refer to this model as a spatial generalized linear model (GLM). It is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution. Since our outcome is binary we will be using a Bernoulli distribution as our target distribution, therefore *logit* will be used as a link function. Forms of the link (10) and mean (11) functions are shown below. For the Bernoulli distribution the interpretation of mean μ is the probability p of y taking the value 1.

$$g(p) = \boldsymbol{x(s)^T\beta} + \omega(\boldsymbol{s}) = \ln\left(\frac{p}{1-p}\right) \tag{10}$$

$$p = \frac{\exp(\boldsymbol{x(s)^T\beta} + \omega(\boldsymbol{s}))}{1+\exp(\boldsymbol{x(s)^T\beta} + \omega(\boldsymbol{s}))} = \frac{1}{1+\exp(-\boldsymbol{x(s)^T\beta} - \omega(\boldsymbol{s}))} \tag{11}$$

In equation 9, $\boldsymbol{x(s)}$ is a $p\times 1$ vector of spatially referenced predictors, including an intercept and spatially referenced covariates. It is associated with a column vector of regression coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_{p-1})$. The space varying impact is captured by $\omega(\boldsymbol{s})$. With any collection of n locations, say $S = \{\boldsymbol{s_1}, ..., \boldsymbol{s_n}\}$, we assume that the elements of $\boldsymbol{\omega}(S)$

are capturing the effect of unmeasured or unobserved mechanisms with spatial pattern and they come from independent Gaussian processes such as $\omega(\boldsymbol{s}) \sim GP(0, C(\boldsymbol{s}, \boldsymbol{s'}; \boldsymbol{\theta}))$. $C(\boldsymbol{s}, \boldsymbol{s'}; \boldsymbol{\theta})$ models the covariance between sites $\boldsymbol{s}$ and $\boldsymbol{s'}$. The process realizations are collected into an $n \times 1$ vector $\boldsymbol{\omega}(\boldsymbol{S}) = (\omega(\boldsymbol{s_1}), \omega(\boldsymbol{s_2}), ..., \omega(\boldsymbol{s_n}))$ which follows a multivariate normal distribution with mean $\mathbf{0}$ and variance $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is the $n \times n$ covariance matrix with (i, j)-th element given by $C(\boldsymbol{s_i}, \boldsymbol{s_j}; \boldsymbol{\theta})$. We further specify $C(\boldsymbol{s}, \boldsymbol{s'}; \boldsymbol{\theta}) = \sigma^2 \rho(\boldsymbol{s}, \boldsymbol{s'}; \phi)$, where $\rho(\cdot, \cdot; \phi)$ is a correlation function and $\boldsymbol{\theta} = \{\phi, \sigma^2\}$ includes parameters quantifying rate of correlation decay, ϕ, and σ^2 is the variance of the spatial component, $Var(\boldsymbol{\omega}) = \sigma^2$ [4], [10]. Package *spBayes* [9] offers four different correlation functions. Here we introduce the exponential correlation function, hence

$$\rho(\boldsymbol{s}, \boldsymbol{s'}; \phi) = exp(-\phi ||\boldsymbol{s} - \boldsymbol{s'}||), \tag{12}$$

where $||\boldsymbol{s} - \boldsymbol{s'}||$ is the Euclidean distance between the two locations $\boldsymbol{s}$ and $\boldsymbol{s'}$.

As we can see in 9, there is no pure error term $\epsilon(\boldsymbol{s})$. According to [12], adding the error term is not sensible since $\omega(\boldsymbol{s})$ is not residual and $\omega(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$ would not be residual either. In this case, the white noise term is replaced by the stochastic mechanism defined by the joint distribution function f,

$$f(y(\boldsymbol{s_1}), ..., y(\boldsymbol{s_n}) | \boldsymbol{\beta}, \sigma^2, \phi) = \int \left[\prod_{i=1}^{n} f(y(\boldsymbol{s_i}) | \boldsymbol{\beta}, \omega(\boldsymbol{s_i})) \right] p(\boldsymbol{\omega} | \sigma^2, \phi) d\boldsymbol{\omega}. \tag{13}$$

4.2 Predictive Process Model

Implementing Markov Chain Monte Carlo algorithms requires repeated evaluation of various full conditional density functions. Particularly, for hierarchical spatial models it requires evaluation of the likelihood, joint or conditional densities arising under the Gaussian process. To obtain such a computation, MCMC algorithms have to work with quadratic forms involving the inverse of the covariance matrix and also its determinant. These computations are very time and memory demanding and for a large number of locations n, it is nearly impossible to obtain the results for a regular user with no special hardware.

For our computations we used PC with Intel Xeon CPU with 3.70GHz and 16GB RAM. In the literature, they refer to these type of problems as *"the big n problems"* [1]. In our particular case, if we would like to fit the model for all our available locations n, we would have to be able to perform computations with matrices of size $213,887 \times 213,887$. Therefore, in this subsection we will describe how to reduce the rank of our model and make the computations possible.

This model is called a predictive process model and was introduced by Banerjee, Gelfand, Finley and Sang [3]. They propose to use a set of *knots*, which may but not need be a subset of recorded locations S. The main requirement is that the number of knots, m, must be much smaller than the total number of locations n, hence $m \ll n$. From the previous section we know that $\omega(\boldsymbol{s}) \sim GP(0, C(\cdot;\boldsymbol{\theta}))$ and let $\boldsymbol{\omega^*}$ be a realization of $\omega(\boldsymbol{s})$ over the set of knots $S^* = \{\boldsymbol{s_1^*}, ..., \boldsymbol{s_m^*}\}$, then $\boldsymbol{\omega^*} \sim MVN(\mathbf{0}, C^*(\boldsymbol{\theta}))$, where $C^*(\boldsymbol{\theta})$ is the associated covariance matrix between sites $\boldsymbol{s_i^*}$ and $\boldsymbol{s_j^*}$. The single site interpolation at site $\boldsymbol{s_0}$ is given by

$$\tilde{\omega}(\boldsymbol{s}) = E[\omega(\boldsymbol{s_0})|\boldsymbol{\omega^*}] = \boldsymbol{c^T}(\boldsymbol{s_0;\theta})C^{*-1}(\boldsymbol{\theta})\boldsymbol{\omega^*}, \tag{14}$$

where $\boldsymbol{c(s_0;\theta)}$ is an $m \times 1$ matrix with entries $\{C(\boldsymbol{s_0, s_i^*;\theta})\}_{i=1}^m$. This idea, based on "kriging" (a method of interpolation for which the interpolated values are modeled by a Gaussian process), defines a spatial process $\tilde{\omega}(\boldsymbol{s}) \sim GP(0, \tilde{C}(\boldsymbol{s, s';\theta}))$ with a covariance function

$$\tilde{C}(\boldsymbol{s, s';\theta}) = \boldsymbol{c^T}(\boldsymbol{s;\theta})C^{*-1}(\boldsymbol{\theta})\boldsymbol{c(s';\theta)}, \tag{15}$$

where $\boldsymbol{c(s;\theta)}$ is an $m \times 1$ matrix with entries $\{C(\boldsymbol{s, s_i^*;\theta})\}_{i=1}^m$. The realizations of $\tilde{\omega}(\boldsymbol{s})$ are precisely the kriged predictions conditional upon a realization of $\omega(s)$ over S^* [1]. Therefore, to obtain the predictive process model we replace the spatial effect $\omega(\boldsymbol{s})$ in equation 9 by the predictive process $\tilde{\omega}(\boldsymbol{s})$, thus

$$g(E[y(\boldsymbol{s})]) = \boldsymbol{x(s)^T\beta} + \tilde{\omega}(\boldsymbol{s}). \tag{16}$$

4.3 Model Specifications and Candidate Models

It the following section we are going to introduce three different candidate models that we are going to work with and their individual model specifications.

For the second and the third model we assign prior distributions to the model parameters to complete Bayesian model specifications. As we mentioned before in 3.3, for modeling purposes, we use R packages *MCMCpack* and *spBayes*, which allow us to fit univariate Bayesian generalized linear non-spatial and spatial regression models with the option to upgrade to a predictive process model with a specified set of knots. The underlying code for the Markov Chain Monte Carlo algorithm is written in C++, which makes computations highly effective for larger sets of observations with *"big n number"* of locations than other available R packages, such as WinBugs, OpenBugs or JAGS.

Natural candidates for priors for $\boldsymbol{\beta}$ parameters have a flat Normal distribution with mean $\mu_{\beta_i} = 0$ and high variance $\sigma^2_{\beta_i}$ for $i = 1, 2, ..., p$. Hence, $\boldsymbol{\beta}$ follows a Multivariate Normal distribution MVN$(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$. The spatial variance σ^2 of the spatial components $\omega(\boldsymbol{s})$ is assumed to follow the Inverse Gamma distribution, $\sigma^2 \sim IG(a_{\sigma^2}, b_{\sigma^2})$. The decay parameter ϕ in the correlation function is assumed to follow a Uniform distribution, $U(a_\phi, b_\phi)$, where a_ϕ and b_ϕ are specified to effectively support over the geographic range of the study area [4], [9].

4.3.1 Generalized Linear Model

First, we would like to fit the simple non-spatial generalized model (GLM). As was mentioned before in section 4.1, we use GLM since our response variable is binary, 0 or 1, and GLM allows us to model other than Normal distributions, in this case a Bernoulli distribution. Since this model is non-spatial we rewrite our link and mean function equations

10 and 11 as:

$$\eta = \boldsymbol{x^T\beta} = \ln\left(\frac{p}{1-p}\right), \tag{17}$$

$$p = \frac{\exp(\boldsymbol{x^T\beta})}{1+\exp(\boldsymbol{x^T\beta})} = \frac{1}{1+\exp(\boldsymbol{-x^T\beta})}, \tag{18}$$

where $\eta = g(E[y])$ is often called the linear predictor and it is related to the expected value (for Bernoulli distribution probability p) of the data via the link function. The coefficients $\boldsymbol{\beta}$ are then derived as maximum likelihood estimates [13].

In R, we can obtain estimates for $\boldsymbol{\beta}$ coefficients using the function `glm(formula, family="binomial", data)`, where `formula` is in the form response $\sim$ covariates. We assessed the model fit performance base on Akaike information criterion (AIC).

AIC is a measure of a relative quality of statistical model. The value of AIC is computed as

$$AIC = 2k - 2\ln(\hat{L}), \tag{19}$$

where k is a number of estimated parameters and $\hat{L}$ is the maximized value of the likelihood function for the model. The final candidate model is then the one with the minimum AIC value [14].

4.3.2 Non-spatial Bayesian Model

The second candidate model uses the Bayesian model structure but in this case we omit the spatial effect. Thus $\omega(\boldsymbol{s}) = 0$ and we can write equation 9 as

$$g(E[y]) = \boldsymbol{x^T\beta}. \tag{20}$$

This model is essentially the same as the previous candidate model, the simple Generalized linear model, since it also uses the logit function as a link function between the linear predictor and the Bernoulli distribution function. The difference here is that to obtain the regression coefficient estimates $\hat{\boldsymbol{\beta}}$ we will not use maximum likelihood but instead we will use Bayesian methods. Therefore, we obtain the posterior distribution as an approximation by the Markov Chain Monte Carlo method. To use MCMC we need to specify *starting* and *prior* values for $\boldsymbol{\beta}$. For the *prior* of $\boldsymbol{\beta}$ we specify flat prior as a Normal distribution with mean 0 and high variance denoted as `Inf` in R. We could also use a specific number, for example, $10,000$. We input 0 as a starting value for each coefficient β_i. We could use package *spBayes* even for a non-spatial model, nevertheless we would like to use a different method for each candidate model, thus we mainly provide results by *MCMCpack*.

4.3.3 Bayesian Hierarchical Model with Spatial Effect

The third model is the main focus of our paper and was already introduced in section 4.1 by equation 9. We will use the same approach as for the previous Bayes model but this time we will add the spatial effect $\omega(\boldsymbol{s})$ to help us account for spatial dependences.

For this model we again have to specify priors for all the parameters. Since regression coefficients are the same as in the previous case we use flat prior for $\boldsymbol{\beta}$, this time using tag `beta.flat` in *spBayes* package. Next we need to introduce support for the exponential decay parameter ϕ. We use $Unif(0.00045, 3)$, which corresponds to support for an effective spatial range between approximately $\sim 1m$ and $7,000m$. We chose $7km$ since it is approximately half of the longest distance between two locations in our data set for individual moose. Recall that since the exponential correlation function is $\rho(\boldsymbol{s}, \boldsymbol{s}^*; \phi) = exp(-\phi||\boldsymbol{s} - \boldsymbol{s}^*||)$, then the effective spatial range is the distance at which the correlation drops to 0.05 [5]. For the spatial variance parameter σ^2 we choose shape and rate as $a_{\sigma^2} = b_{\sigma^2} = 2$, which gives us $IG(2, 2)$. We use this setting since we expect the variance to be close to 2 and the mean of an Inverse Gamma distribution is equal to $E[X] = b_{\sigma^2}/(a_{\sigma^2} - 1)$ for $a_{\sigma^2} > 1$.

Now, when we specified all the parameters for the prior distributions, we can write the posterior distribution $p(\theta|y)$ in a similar fashion as in equation 8 such as

$$\begin{aligned} p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \prod_{i=1}^{n} & f(y(\boldsymbol{s_i})|\boldsymbol{x}(\boldsymbol{s_i})^T\boldsymbol{\beta} + \omega(\boldsymbol{s_i})) \\ & \times N(\boldsymbol{\omega}|\mathbf{0}, \boldsymbol{\Sigma}) \times N(\boldsymbol{\beta}|\mathbf{0}, \boldsymbol{\Sigma_\beta}) \\ & \times IG(\sigma^2|a_{\sigma^2}, b_{\sigma^2}) \times U(\phi|a_\phi, b_\phi), \end{aligned}$$

where $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\omega}, \sigma^2, \phi\}$ is the set of all parameters and $\boldsymbol{y} = (y(\boldsymbol{s_1}), y(\boldsymbol{s_2}), ..., y(\boldsymbol{s_n}))$ is the vector of dependent variables. In figure 3 we can see a Bayesian network for our model. Bayesian networks are drawings that depict probability distributions and the relationships among the parameters graphically [2]. The deterministic relationship between the model and covariates is shown with dashed arrow. Relationships between random variables are represented by solid arrows.

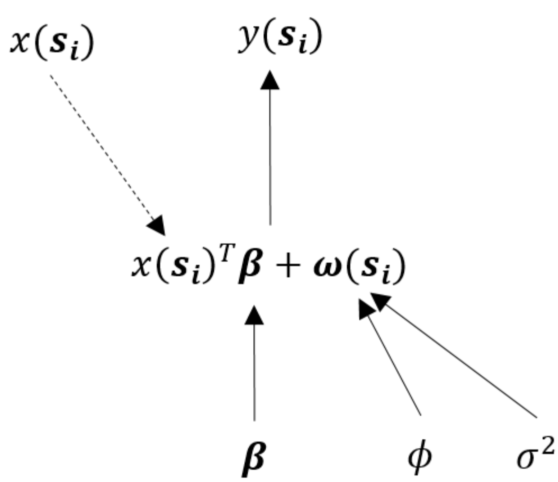

Figure 3: Bayesian network.

As we mentioned in section 3.3, we also have to specify the *starting* values for our parameters. We choose these values for parameters σ^2 and ϕ as the most informative educated assumption based on our knowledge of the data. We use $\sigma^2 = 2$ and $\phi = 3/50$, which should correspond with effective spatial range of $50m$. For the coefficients $\boldsymbol{\beta}$ we obtain the starting parameters as an estimate of the coefficients from the simple Generalized linear model fit. Initial values should not affect final results since the MCMC algorithm will eventually converge to the "true" value, however a good choice of starting parameters will speed up the convergence.

We quantify how well the Bayesian models fit the data by the deviance information criterion (DIC). It is a hierarchical modeling generalization of the Akaike information criterion (AIC), which is define as

$$DIC = \bar{D} + p_D, \tag{21}$$

where $\bar{D}$ is the posterior mean of the deviance (a measure of model goodness of fit) and p_D is the effective number of parameters (a penalty for model complexity). DIC with lower values indicates a better model fit. For the non-spatial models with weak priors DIC≈AIC [14].

5 Results

In this section we would like to show and discuss the results of our proposed models. We split this section into two parts where we look first at the results of the non-spatial models and then at the results given by the model with a spatial varying component. We approached it this way because of the computational challenge that arises from the spatial model, where we can not use all the available observations. Nevertheless, we provided comparison between both non-spatial and spatial models in subsection 5.2.

5.1 Non-spatial Models

The results for non-spatial models were obtained by using all the observations for each moose separately and also for all ten moose together. We mainly performed the computations by R package *MCMCpack*, but we also obtained the results from the non-spatial version offered in the *spBayes* package. As we mentioned before in 4.3.2, these non-spatial models (GLM and Bayes) do not have a different model framework, but the difference comes from how the estimates of the regression coefficients β_i are obtained. Therefore, we would like to see if one or the other gives us better results and if there are any positives or negatives for choosing the best approach for the non-spatial generalized linear models. To measure the quality of the model estimates we used Mean Square Error (MSE) computed as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2. \tag{22}$$

For the non-spatial Bayesian models, the MCMC chains were run for $30,000$ iterations each and the first 75% of outcomes were burned. The computations with *MCMCpack* were faster than with *spBayes*, however the *spBayes* outcome includes more information such as deviance and a prepared covariate matrix $\boldsymbol{X}$ for future estimates of $\hat{\boldsymbol{Y}}$. We omitted the results from *spBayes* in this section; however, we have let it run for all the moose and based on the deviance we obtained a Deviance information criterion (DIC) for all moose and as was mentioned in section 4.3.3 DIC was nearly identical with AIC given by the GLM model. Since the outcome estimates for β_i were in the form of *logit*, they represent an increase in ln odds. Therefore, we computed the odds ratios as $\exp(\beta_i)$ to obtain the odds ratio associated with a one-unit increase in the exposure. We also computed the 95% confidence interval for the odds ratio, which shows insignificance if the interval includes number 1.

In this part of the section we show four tables for two different moose. Table 4 and 5 for moose with ID 31166 and table 6 and 7 for ID 31179. Tables for other moose can be found in Appendix I - Tables for the non-spatial models using all observations. These two moose are specific by a high frequency of Hydro "PRESENT" values, 743 and 859, respectively. However, they are also representatives of two groups with different numbers of observations. Recall table 1, the frequency of observations after removing the outliers from vegetation type covariate is $28,453$ and $14,390$, respectively. We may see that both methods of estimating the β_i parameters returned similar values, as was expected. This is true for most of the moose; nevertheless, there are exceptions for some of the covariates. We will return to this issue in the future.

Let us now look at the individual covariates and their results. Starting with the Collar temperature, based on p-values the estimate is significant for most of the moose. We can see values slightly higher than 0 as a β_1 estimate, therefore odds ratios are slightly above 1. This suggests that with the rising temperature the moose tend to seek water. Still, the maximum value among all the moose is $\beta_1 = 0.052 \rightarrow OR = 1.053$, which does not show a big increment due to the change in temperature.

Secondly, we have the vegetation type covariate `CLASS_CODE_NEW` split into three regressors. We use as a reference the group number 2, which includes vegetation with a moderate amount of water. We may observe in tables 4 and 6 differences in values of the estimates and also in the significance of individual regressors. However, the common pattern among all the moose is that the regression coefficient for group 1 ("Open water")

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.916	0.181	-16.143	$<$0.001	0.054	0.038	0.077
Collar Temperature	0.043	0.007	6.199	$<$0.001	1.044	1.030	1.059
Open Water	1.262	0.099	12.737	$<$0.001	3.532	2.909	4.289
Mixedwood Forest	-3.340	0.141	-23.629	$<$0.001	0.035	0.027	0.047
Low Density Forest	-4.175	0.385	-10.836	$<$0.001	0.015	0.007	0.033
Time Day/Night	0.464	0.113	4.096	$<$0.001	1.590	1.274	1.985
04/15 - 05/31	-0.587	0.114	-5.155	$<$0.001	0.556	0.445	0.695
08/01 - 10/31	-1.697	0.124	-13.686	$<$0.001	0.183	0.144	0.234

Table 4: Generalized linear model, Moose ID: 31166

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.918	0.193	-15.127	$<$0.001	0.054	0.037	0.079
Collar Temperature	0.044	0.008	5.730	$<$0.001	1.045	1.029	1.060
Open Water	1.266	0.101	12.483	$<$0.001	3.547	2.908	4.327
Mixedwood Forest	-3.352	0.145	-23.089	$<$0.001	0.035	0.026	0.047
Low Density Forest	-4.240	0.404	-10.490	$<$0.001	0.014	0.007	0.032
Time Day/Night	0.455	0.120	3.783	$<$0.001	1.575	1.245	1.994
04/15 - 05/31	-0.589	0.117	-5.027	$<$0.001	0.555	0.441	0.698
08/01 - 10/31	-1.703	0.123	-13.870	$<$0.001	0.182	0.143	0.232

Table 5: Bayesian Non-spatial model (MCMCpack), Moose ID: 31166

	Estimate	**SE**	**z value**	**p-value**	**Odds ratio**	**L 0.95 CI**	**U 0.95 CI**
Intercept	-0.529	0.328	-1.614	0.106	0.589	0.310	1.120
Collar Temperature	0.029	0.008	3.895	<0.001	1.030	1.015	1.045
Open Water	-0.078	0.290	-0.270	0.787	0.925	0.524	1.632
Mixedwood Forest	-2.804	0.296	-9.462	<0.001	0.061	0.034	0.108
Low Density Forest	-3.716	0.295	-12.592	<0.001	0.024	0.014	0.043
Time Day/Night	-0.063	0.093	-0.671	0.502	0.939	0.782	1.128
04/15 - 05/31	-0.775	0.140	-5.540	<0.001	0.461	0.350	0.606
08/01 - 10/31	-0.789	0.104	-7.591	<0.001	0.454	0.370	0.557

Table 6: Generalized linear model, Moose ID: 31179

	Estimate	**SE**	**z value**	**p-value**	**Odds ratio**	**L 0.95 CI**	**U 0.95 CI**
Intercept	-0.516	0.333	-1.550	0.121	0.597	0.311	1.146
Collar Temperature	0.029	0.008	3.735	<0.001	1.029	1.014	1.045
Open Water	-0.083	0.287	-0.288	0.773	0.921	0.525	1.616
Mixedwood Forest	-2.820	0.295	-9.567	<0.001	0.060	0.033	0.106
Low Density Forest	-3.719	0.300	-12.399	<0.001	0.024	0.013	0.044
Time Day/Night	-0.068	0.093	-0.734	0.463	0.934	0.779	1.120
04/15 - 05/31	-0.768	0.145	-5.284	<0.001	0.464	0.349	0.617
08/01 - 10/31	-0.799	0.103	-7.736	<0.001	0.450	0.367	0.551

Table 7: Bayesian Non-spatial model (MCMCpack), Moose ID: 31179

has a positive estimate and for group 3 ("Mixedwood forest") and 4 ("Lower density forest") it is negative. This distribution is expected since the odds ratios show that increasing exposure to the location with a water feature increases the odds of the moose being in the water. Another mentionable outcome is a lower value for group number 4; therefore, the odds tend to decrease for the lower density forests. We are able to see this pattern for all the moose where all the class code regressors are significant.

The third covariate is `Time_NEW`. Since this variable has only values 0 or 1, a positive estimate suggests a higher chance of moose seeking water during the daylight, negative after the sunset. In tables 4 and 5 we have the positive and in 6 and 7 we have the negative value for `Time Day/Night`. But for the moose with ID 31179 this estimate is insignificant. Yet, if we look at all the moose, most of them have a negative significant value for the time coefficient, thus this covariate does not support our expectation that the moose are rather close to the water during the day.

Last but not least, we have the covariate `Date_NEW`, which splits the whole observation period into three parts. We use as a reference the part number 2, which includes the season of the highest temperatures. Therefore, our assumption is the other two parts will have negative regressor coefficients. If we look at our two moose, we can observe that our expectation holds. The same or similar results can be seen also for the other moose. Hence, summer months have a positive effect on moose occurrence near the water. Moreover, in most of the cases the coefficient for part 1 (April, May) is lower; thus, we could expect a lower frequency of Hydro "PRESENT" for these months. Of course, to show this we would need more data, since group 1 has the lowest density, because our data set includes only half of month of April.

In terms of comparison how well each model fits the data, we computed MSE for three different outcomes. The Generalized linear model gives us fitted values as a probability vector $\hat{\boldsymbol{p}}$ of size $n \times 1$. For the non-spatial Bayesian model, we first need to compute these probabilities using equation 11. There is of course a difference for the Bayesian outcome. Since we burned 75% of the MCMC chain, we now have a posterior sample for

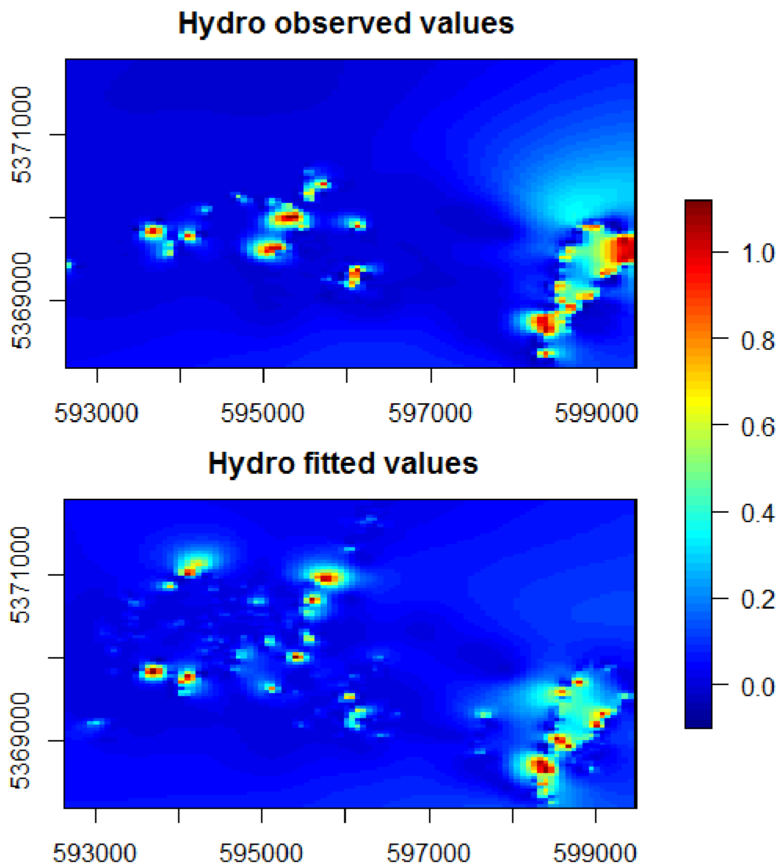

Figure 4: Observed vs. fitted values given by GLM, moose ID: 31179

each parameter with $7,500$ values. Thus, for the Bayesian model we have a matrix of probabilities $\hat{\boldsymbol{P}}$ of size $n \times 7,500$. We then compute the fitted binary output using the *binom* function in R with n number of observations, 1 trial for each observation and a probability given by $\hat{\boldsymbol{p}}$ or $\hat{\boldsymbol{P}}$. Since our result matrix for the Bayesian outcome is still $n \times 7,500$, we count the number of 1's for each observation (number of 1's in each row) and if the sum is greater than half the size of the whole row ($> 3,750$), then we assume that the fitted value for this location is equal to 1, otherwise 0. By this computation we obtain a vector $\hat{y}$ of size $n \times 1$ with only 0's and 1's. We then compare the binomial fitted values with the observed values for the parameter Hydro to compute MSEs. We call them `GLM Binomial` and `Bayes Binomial`. We also compute MSEs for the observed values of Hydro and the probability matrices $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{P}}$ and call them `GLM Probability` and `Bayes`

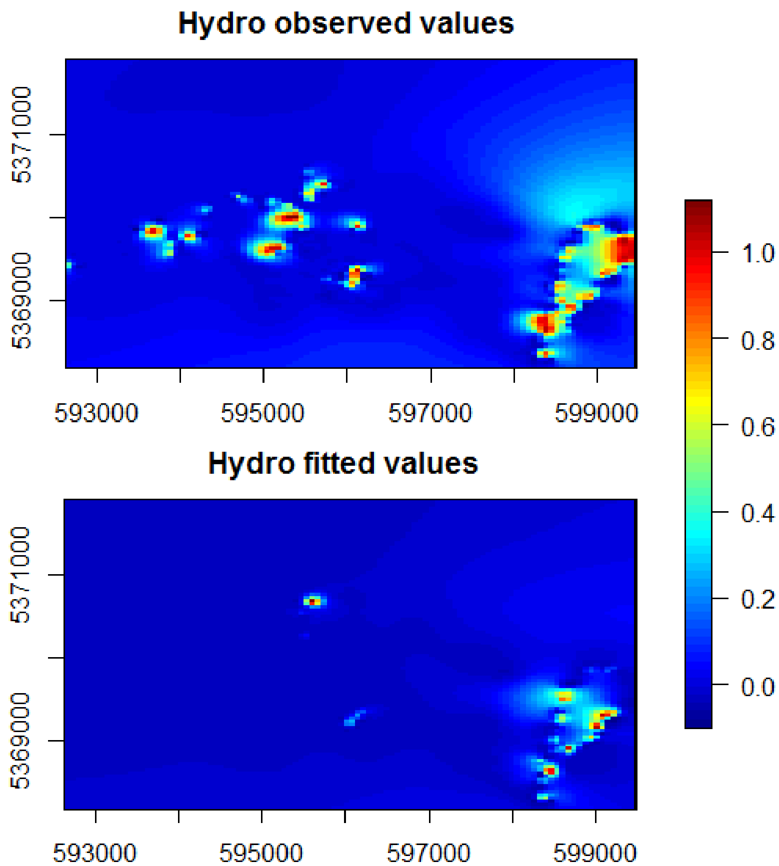

Figure 5: Observed vs. fitted values by Bayes non-spatial model, moose ID: 31179

`Probability`. As a third comparison we rescale $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{P}}$ such that we want the location with the highest probability to have value 1. The MSEs for the transform probabilities are called `GLM Probability scaled` and `Bayes Probability scaled`. We summarize the comparison by MSE in table 8. By looking at the table 8 we can see that the values for MSEs computed using the probability matrices gave us the same results for both models or slightly higher numbers for the Bayes model. But analyzing the results given by the "binomial" outcome, with exception of MSE for all the moose together, each moose shows better fit for the Bayesian model. However, this is connected with under-fitting of the Bayesian model. For each moose the number of the fitted locations with Hydro "PRESENT" is lower than the observed count. On the other side, GLM returns a similar amount of locations in the water, but usually higher. Therefore, there is a higher amount of error for GLM.

We can see this very well in figures 4 and 5, where we show a comparison between the observed and the fitted values based on the binomial output described above for the moose with ID 31179. Figure 4 shows a higher number of fitted locations with value 1, but it is noticeable that a lot of locations do not match the observed values.

	ID: 31166	**ID: 31168**	**ID: 31169**	**ID: 31170**
GLM Binomial	0.0381	0.0275	0.0266	0.0724
Bayes Binomial	0.0256	0.0152	0.0149	0.0444
GLM Probability	0.0194	0.0137	0.0140	0.0362
Bayes Probability	0.0194	0.0137	0.0140	0.0362
GLM Probability scaled	0.0212	0.0139	0.0163	0.0394
Bayes Probability scaled	0.0212	0.0139	0.0163	0.0395
	ID: 31172	**ID: 31174**	**ID: 31178**	**ID: 31179**
GLM Binomial	0.0284	0.0115	0.0055	0.0833
Bayes Binomial	0.0140	0.0058	0.0032	0.0600
GLM Probability	0.0133	0.0055	0.0030	0.0427
Bayes Probability	0.0133	0.0055	0.0030	0.0427
GLM Probability scaled	0.0200	0.0055	0.0031	0.0483
Bayes Probability scaled	0.0197	0.0055	0.0031	0.0485
	ID: 31182	**ID: 31184**	**ALL moose**	
GLM Binomial	0.0021	0.0133	0.0296	
Bayes Binomial	0.0013	0.0073	0.0299	
GLM Probability	0.0011	0.0067	0.0149	
Bayes Probability	0.0011	0.0067	0.0149	
GLM Probability scaled	0.0012	0.0068	0.0161	
Bayes Probability scaled	0.0012	0.0068	0.0161	

Table 8: Mean Square Error for all moose.

We mentioned that the estimates of regression coefficients for the GLM and the Bayes model are similar in most of the cases, especially if those estimates are significant. But there exist exceptions. A good example of this situation is moose with ID 31182. The GLM estimates for class code groups 3 and 4 are -17.56 and -17.49, respectively. In case of the Bayes model those estimates are -374.85 and -312.23, respectively. We can observe a huge difference between the outcomes of these two models. This is caused by the inability of the MCMC chain to converge during the $30,000$ iterations, since this moose has a low frequency of Hydro "PRESENT" values (only 17 in $n = 14,390$). We tried to run a chain for this moose with a higher amount of iterations, but it did not improve the final estimates or did not make them close to the GLM result. Nevertheless, in all the cases the regressors were not significant.

As a last topic in this subsection, we would like to look at the results for all the ten moose together. Tables 9 and 10 represent the GLM and the Bayes model outputs. The Bayesian computation required approximately a half hour to finish $30,000$ iterations. We may see that both tables are nearly identical and all the regressors are significant at the level $\alpha = 0.05$. We may notice that the pattern described above for the individual moose holds pretty well.

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.482	0.074	-46.988	<0.001	0.031	0.027	0.036
Collar Temperature	0.024	0.003	8.106	<0.001	1.024	1.018	1.030
Open Water	3.266	0.046	70.642	<0.001	26.208	23.937	28.693
Mixedwood Forest	-1.656	0.051	-32.334	<0.001	0.191	0.173	0.211
Low Density Forest	-1.222	0.061	-20.190	<0.001	0.295	0.262	0.332
Time Day/Night	-0.104	0.041	-2.527	0.012	0.901	0.831	0.977
04/15 - 05/31	-0.901	0.056	-16.036	<0.001	0.406	0.364	0.453
08/01 - 10/31	-0.594	0.041	-14.310	<0.001	0.552	0.509	0.599

Table 9: Generalized linear model for all moose.

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.483	0.076	-45.773	<0.001	0.031	0.026	0.036
Collar Temperature	0.024	0.003	7.997	<0.001	1.024	1.018	1.030
Open Water	3.267	0.047	68.791	<0.001	26.240	23.908	28.800
Mixedwood Forest	-1.658	0.052	-31.983	<0.001	0.190	0.172	0.211
Low Density Forest	-1.222	0.061	-19.868	<0.001	0.295	0.261	0.333
Time Day/Night	-0.102	0.041	-2.463	0.014	0.903	0.833	0.979
04/15 - 05/31	-0.902	0.056	-15.971	<0.001	0.406	0.363	0.453
08/01 - 10/31	-0.594	0.039	-15.069	<0.001	0.552	0.511	0.597

Table 10: Non-spatial Bayesian model for all moose.

5.2 Spatial Model

In the previous section we showed the results of the non-spatial models. Now we summarize the outcome of the model with the spatial effect. We used the R package *spBayes*, since it allows us to use the predictive process to reduce the dimension of the covariance matrix used for calculations as described in section 4.1. We specified the number of knots m to be 10% of the number of observations n, thus $m = 0.1 * n$. The knots are then selected by method of k-means clustering. Distribution of the selected locations and the knots is represented in figure 6. However, even with this dimension reduction the computation process is very challenging for memory and speed. We can not even start the MCMC chain for the moose with $28,000+$ observations, since we are unable to allocate the covariance vectors in the memory. An estimated upper bound for the number of observations based on trial runs is $15,000$. But then there arises another issue with the speed of the computation. Based on the tests, one MCMC chain run with the $30,000$ iterations for the $n = 15,000$ observations would take approximately 30 days with the available hardware. Therefore, we decided to run each chain for $n = 5,000$ randomly selected locations for each moose. Setting the acceptance rate to 30% within the adaptive MCMC Metropolis-within-Gibbs algorithm (see section 3.3 and [11]), the average time of the one MCMC chain was then ~ 33 hours.

For the purposes of comparison we will again use the moose with ID 31166 so we can analyze the results also with the previous subsection 5.1. We will compare the outcomes of the Generalized linear model (Table 11) and the spatial Bayes model (Table 12), but in this case we will also use the results obtained by running the non-spatial Bayes model. We will first focus on the regression coefficients β_i. We might see that there are differences between the GLM and the Bayes model output. However, the change is only in the amplitude and there is no big difference which would modify the signs of the regressors. Change of sign happens only in the case of insignificant results. Therefore the "direction" of odds remains unchanged. By looking at tables 11 and 12 and other moose's tables available in Appendix II - Tables for the spatial model using selected observations, we can observe similar results and patterns as for the non-spatial models in previous subsection 5.1. The collar temperature shows again a small increase in occurrence of the moose near

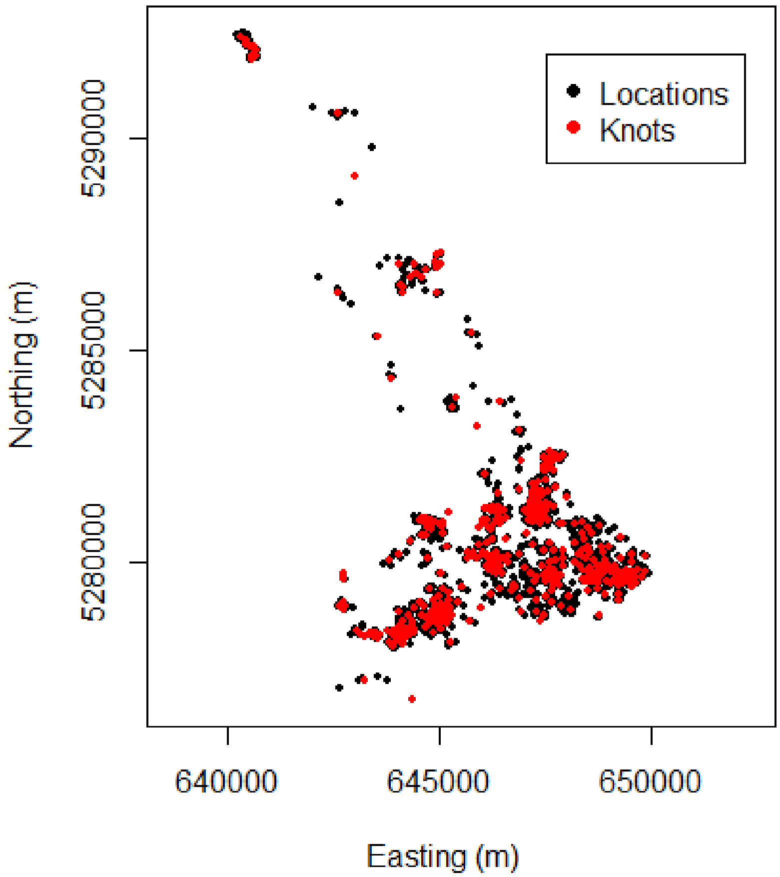

Figure 6: Distribution of location and knots for moose ID 31166 with $n = 5,000$ and $m = 500$.

the water with a rising temperature. Still, this coefficient contributes just a small portion to the total variability and for some of the moose it is not significant. The highest numbers are associated with the class code of the vegetation type in group 1. This outcome was again expected, since the group 1 represents the location in a water feature. Therefore, we should be alert in the case of moose with ID 31172, since the estimates for `Open Water` are -12.737 and -336.800 for the GLM and the Bayes model, respectively. In this situation the estimates are both insignificant and were caused by the random selection of our n observations, since a minimum of points with the class code 1 is included in the sample for this particular moose. The other two vegetation type groups have again the odds ratio lower than 1. Usually the odds are lower for the lower density forests (class code 4), but we may also see the opposite result for the moose with ID 31172. Some of the moose have very low, insignificant values, which are caused, again, by the random sampling from all the observations. Tables 11 and 12 suggest that there is a higher appearance of the moose near the water during the day, but this result is not significant. Also for the other moose the estimated odds ratios are exactly opposite, similar to the case of the non-spatial models. However, the coefficient for the time effect is rarely significant for the Bayesian model, thus it is hard to derive any conclusion. We have the same difficulty with the `Date_NEW` regression coefficient since in most cases the estimates for both the GLM and the Bayes spatial model are insignificant. Still, by comparing all the results, it shows the similar pattern; thus, we expect a lower occurrence of the Hydro "PRESENT" values for the beginning of spring and late summer/start of fall.

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.250	0.435	-7.466	<0.001	0.039	0.017	0.091
Collar Temperature	0.064	0.017	3.763	<0.001	1.066	1.031	1.102
Open Water	1.314	0.237	5.547	<0.001	3.722	2.339	5.922
Mixedwood Forest	-3.452	0.357	-9.655	<0.001	0.032	0.016	0.064
Low Density Forest	-4.407	1.015	-4.342	<0.001	0.012	0.002	0.089
Time Day/Night	0.174	0.266	0.656	0.512	1.191	0.707	2.005
04/15 - 05/31	-0.434	0.277	-1.565	0.117	0.648	0.377	1.115
08/01 - 10/31	-1.573	0.297	-5.292	<0.001	0.207	0.116	0.371

Table 11: Generalize linear model for moose ID 31166 with $n = 5,000$.

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-6.000	0.710	-8.457	<0.001	0.002	0.001	0.010
Collar Temperature	0.076	0.022	3.490	<0.001	1.079	1.034	1.126
Open Water	2.549	0.439	5.800	<0.001	12.792	5.406	30.270
Mixedwood Forest	-2.491	0.582	-4.277	<0.001	0.083	0.026	0.259
Low Density Forest	-4.650	1.319	-3.524	<0.001	0.010	0.001	0.127
Time Day/Night	-0.028	0.320	-0.088	0.930	0.972	0.519	1.820
04/15 - 05/31	0.416	0.420	0.990	0.322	1.515	0.666	3.448
08/01 - 10/31	-0.334	0.586	-0.570	0.569	0.716	0.227	2.258
sigma.sq	6.109	1.998	3.057	0.002			
phi	0.006	0.003	2.266	0.023			

Table 12: Bayesian model with spatial effect for moose ID 31166 with $n = 5,000$.

The spatial variance σ^2 shows the spatial structure in the non-spatial residuals. Higher values indicate a greater spatial structure; hence, the need to include a spatial varying coefficient. In table 12, the spatial variance is equal to 6.109, which suggests that including the spatial parameter could improve model fit. The σ^2 is significant in most of the cases and ranges from ~ 2.00 to ~ 30.00. Furthermore, for example, the moose with ID 31179 has $\sigma^2 = 501.816$, which indicates a great spatial dependence, therefore including the spatial varying coefficient should vigorously improve the model fit performance. We present the observed and fitted values for the moose with ID 31166 in figure 7. In figure 8 and 9 we show the observed and fitted values for the GLM and the Bayes spatial model for moose ID 31179, respectively. The spatial models are not calculated for all the available locations, but we can still observe better fitting performance by comparison with figure 5 for moose ID 31179.

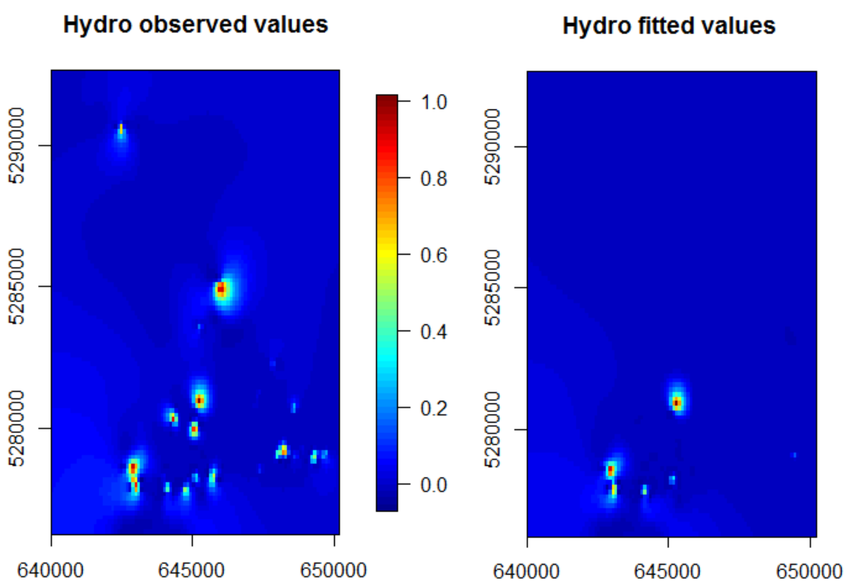

Figure 7: Observed vs. fitted values for Bayesian spatial model for moose ID 31166 with $n = 5,000$.

The effective spatial range ϕ is another parameter that describes the strength of the spatial component. A long spatial range suggests stronger and farther reaching spatial dependence. If we want to convert the spatial range ϕ into meters we use $3/\phi$, since

$|\ln 0.05| \approx 3$. Therefore, from table 12 we have $\phi = 0.006$; thus, the effective spatial range is equal to $500m$. The effective spatial range differs from $\sim 400m$ to $\sim 1000m$. The long spatial ranges suggests that the spatial random effect was capturing a substantial residual spatial structure and the non-spatial models are not appropriate. Only moose 31178 and 31182 have much shorter values, which is connected with the very small amount of the Hydro "PRESENT" values.

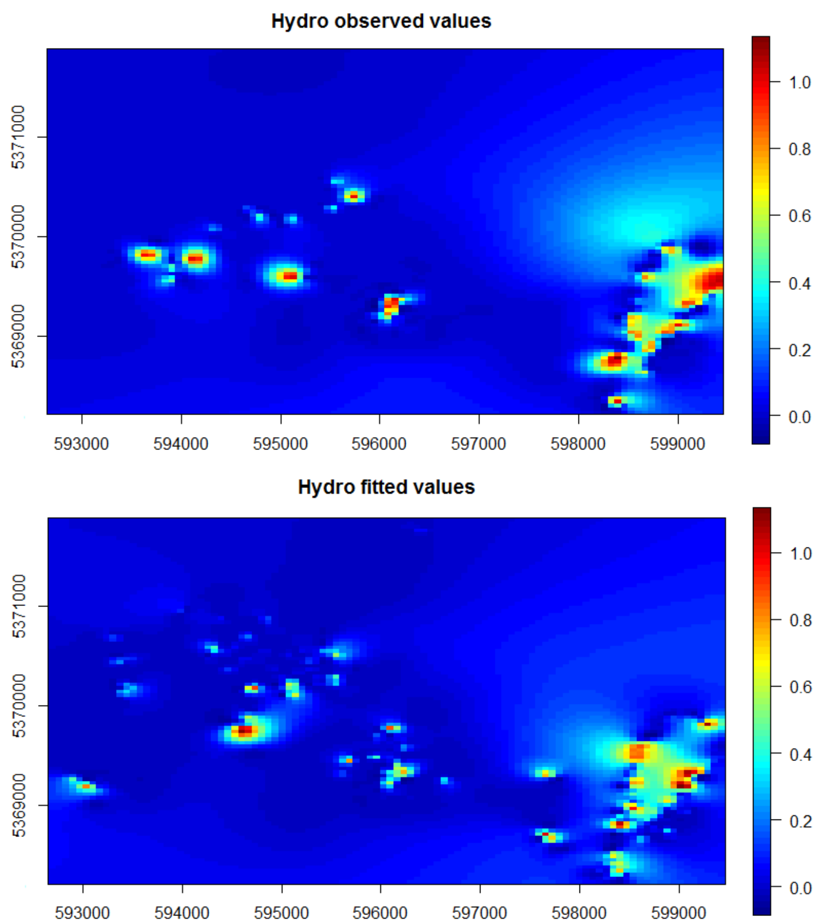

Figure 8: Observed vs. fitted values for Generalized linear model for moose ID 31179 with $n = 5,000$.

The Bayesian hierarchical model performance was assessed using the Deviance information criterion (DIC). We compare the values of DIC with the non-spatial models AIC

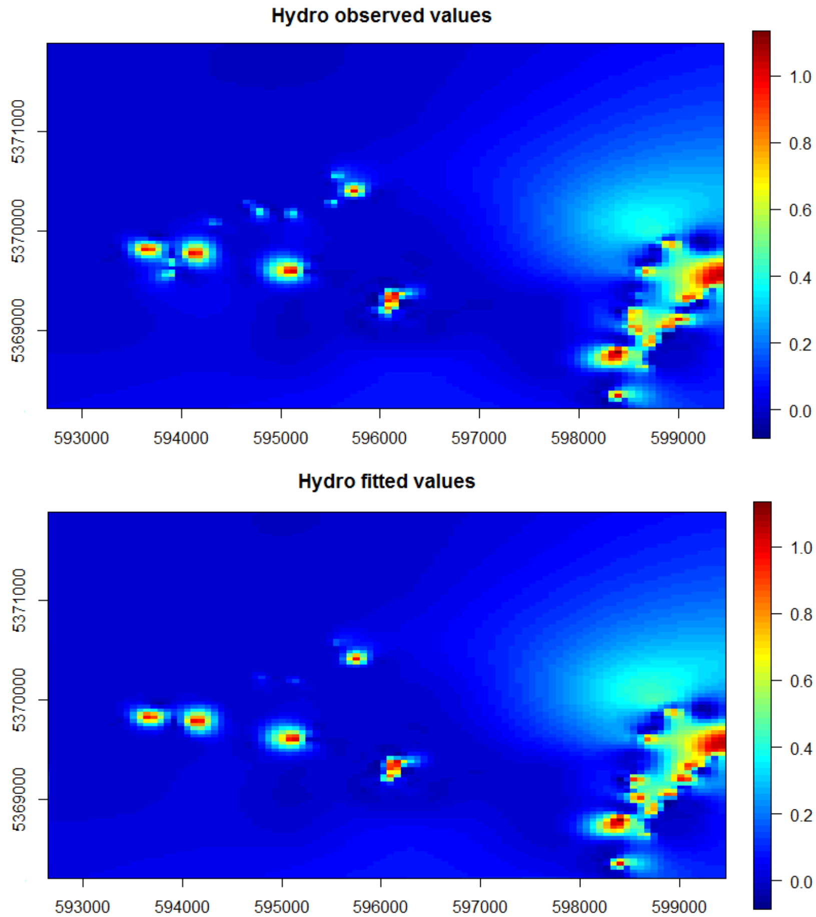

Figure 9: Observed vs. fitted values for Bayesian spatial model for moose ID 31179 with $n = 5,000$.

and DIC. As was mentioned before in 4.3.3, for the non-spatial models AIC≈DIC. We verify this relationship by running both the Generalized linear and the Bayesian non-spatial model. Thus in table 13 we can compare the AIC of the Generalized linear model, the DIC of the Bayesian non-spatial model and the DIC given by the Bayesian hierarchical model with spatially varying coefficient. We can see that the AIC and the DIC of the non-spatial models are very similar and the biggest difference is equal to 5.1 for ID 31182. It is difficult to say what would constitute as a significant difference in DIC, since there is no strict rule. Very roughly, differences of more than 10 might definitely rule out the model with the higher DIC [14]. Thus we can see that there is no substantial difference between the model fit performance for the GLM and the non-spatial Bayes model. However, we can notice significant differences between the non-spatial models and the model with spatially varying coefficient. The difference in the DIC is at least ≈ 100, with exception for 31178 and 31182. These two moose where mentioned before, because we did not see much of an effect of the spatial component from the spatial variance and the spatial range. In both cases, it is given by a small occurrence of the Hydro "PRESENT" in their data. For 31178 it could also be caused by the random sampling from all the observations. We can also observe an opposite result for the moose with ID 31179, where the difference is 1048.5. The data for 31179 showed a strong spatial dependence, therefore including the spatially varying coefficient substantially improved the fitting performance. The same result we expected for ID 31170, but we were unable to derive any value for DIC. We ran the MCMC chain twice for this moose, but we did not receive any better results. This error is probably connected with the fact that during the posterior sampling process of MCMC for σ^2 in one moment the graph shows a huge spike, which probably made it impossible to compute the DIC. We also computed the same MSE statistics as in previous section 5.1. We omit the

	ID: 31166	**ID: 31168**	**ID: 31169**	**ID: 31170**
GLM AIC	710.7	635.7	652.0	1332.3
non-spatial DIC	710.5	633.2	652.1	1333.4
spatial DIC	538.4	332.6	336.8	N/A
	ID: 31172	**ID: 31174**	**ID: 31178**	**ID: 31179**
GLM AIC	584.0	277.7	215.6	1511.2
non-spatial DIC	582.1	276.0	213.4	1513.7
spatial DIC	391.1	189.1	197.3	465.2
	ID: 31182	**ID: 31184**	**All moose**	
GLM AIC	54.5	385.2	1071.0	
non-spatial DIC	49.4	385.8	1071.0	
spatial DIC	48.7	312.2	890.0	

Table 13: Fit statistics.

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.788	0.389	-9.747	<0.001	0.023	0.011	0.048
Collar Temperature	0.038	0.015	2.498	0.012	1.039	1.008	1.071
Open Water	3.408	0.242	14.081	<0.001	30.195	18.791	48.522
Mixedwood Forest	-1.529	0.252	-6.058	<0.001	0.217	0.132	0.356
Low Density Forest	-1.431	0.334	-4.290	<0.001	0.239	0.124	0.460
Time Day/Night	0.001	0.216	0.005	0.996	1.001	0.656	1.528
04/15 - 05/31	-0.939	0.290	-3.241	0.001	0.391	0.222	0.690
08/01 - 10/31	-0.649	0.216	-2.999	0.003	0.523	0.342	0.799

Table 14: Generalize linear model for all moose with $n = 8,000$.

	Estimate	SE	z value	p-value	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-5.629	0.620	-9.085	<0.001	0.004	0.001	0.012
Collar Temperature	0.055	0.018	3.130	0.002	1.057	1.021	1.094
Open Water	3.179	0.429	7.412	<0.001	24.035	10.368	55.718
Mixedwood Forest	-1.788	0.372	-4.809	<0.001	0.167	0.081	0.347
Low Density Forest	-1.975	0.465	-4.246	<0.001	0.139	0.056	0.345
Time Day/Night	0.075	0.243	0.309	0.757	1.078	0.669	1.736
04/15 - 05/31	-0.978	0.391	-2.503	0.012	0.376	0.175	0.809
08/01 - 10/31	-0.170	0.318	-0.536	0.592	0.843	0.452	1.573
sigma.sq	6.584	1.631	4.038	<0.001			
phi	0.002	0.001	3.813	<0.001			

Table 15: Bayesian spatial model for all moose with $n = 8,000$.

results in this paper; nevertheless, for all the moose we received better (lower) values from the Bayes model with a spatially varying coefficient.

Finally, we derived results also for all the moose together. We decided to use a random sample of approximately $8,000$ observations collected together from all 10 moose. We would like the sample from each moose to be proportional to its total number of observations n_i for $i = 1, ..., 10$. Therefore, we used formula $(n_i/N) \cdot 8,000$, where $N = 213,632$ is the total number of observations, to obtain amount of observations for each moose. This gave us a random sample of $8,004$ observations. Table 14 shows the Generalized linear model results. Table 15 then shows the results obtained by the Bayesian model with spatially varying coefficient. Run time of the Bayesian model was 162.64 hours.

With the exceptions of variable `Time Day/Night` for both models and `08/01 - 10/31` for spatial model all estimates are significant. We notice the similar outcomes as we presented for the individual moose. Therefore, we can see slight increase for higher temperature and large impact of the location vegetation type. We did not obtain the same results for other class code groups for both models. However, the spatial model suggests lower odds of moose seeking the water source for a lower density forest, which was also our previous observation based on individual moose and non-spatial model in 5.1. Covariate `Date_NEW` again suggests the higher odds of positive outcome for summer months June and July. If we look at the spatial variance σ^2 in table 15 it shows similar numbers as, for example, moose with ID 31166 or 31184. The value 6.584 is definitely not that low to show no spatial dependence. The lower value is given, as in the cases of individual moose, by small amount of Hydro “PRESENT” values in the random sample. Nevertheless, the amount is not that small and the model was still able to capture some of the residual spatial structure. The effective spatial range $3/\phi$ is equal to approximately $1,500m$, which is the longest spatial range compare to the individual moose. This value was expected and it could be possibly even higher given to the spread of all the location. The spatial range was possibly capturing the spatial dependences for moose which were living close to each other. The spatial model improved the model fitting performance by 181 points (table 13) and also MSE for the binomial response improved from 0.0292 for GLM to 0.0148.

6 Conclusion and Future Work

Starting with describing the data set, our goal was to derive some behavioral statistic about the moose use of water in the Northern Minnesota. We prepared the collected data such that we could use them to fit our models proposed later. We also added or modified some currently present variables to be able to address specific points of interests such as an effect of time of the day or a particular season. Our main goal was to utilize the Bayesian hierarchical model with spatially varying coefficients to fit our data for the variable Hydro. Hydro represents the locations where moose are in the water future mainly to cool down themselves, walk and feed on aquatic plants. We described the Bayesian theorem and the theory behind sampling from the posterior distribution. To be able to derive any samples from the posterior distribution we also needed to describe the Monte Carlo Markov Chain algorithm and R packages *MCMCpack* and *spBayes* we used to obtain these samples. Next, we introduced the general framework of the Generalized Bayesian model and gave explanations for all the parameters used in the model. We also presented the modification with the predictive process we employed to lower computational burdens connected with a large amount of different locations. Last but not least, we described three candidate models: Generalized linear model, Non-spatial Bayesian model and Bayesian hierarchical model with spatially varying coefficients.

Based on the results obtained from these models we were able to estimate the parameter values, significance and odds ratios. We then derived some information about the moose behavior such as there is an evidence that with rising temperature the moose are more likely be near water features. This is also highly associated with the vegetation type where the moose is currently located. In the case of an effect of the time of the day we expected a higher occurrence of the moose in the water feature during the day. However, we could not seen this from our results. Rather, we saw the opposite effect on the moose behavior. The seasonal effect gave us substantial evidence that the summer months such as June and July have a positive effect on the outcome of Hydro. We could also see higher odds for the moose to be seeking water for months at the end of the summer (July) and the beginning of fall (September, October), rather than at the beginning of spring (April, May).

In the case of overall fit performance we found a significant improvement if we use the model with the spatial effect. With the exception of moose with ID 31178 and 31182 we found evidence of substantial spatial dependence among residuals and employing the spatially varying coefficient boosted the model fit performance measured by the Deviance Information Criterion (DIC) and the Mean Square Error (MSE).

In terms of the future work, it would definitely help to have a bigger computational power so we could run the MCMC chains for all the available locations. Also we could possibly run more than one chain for the same moose and obtain the fitted values from the posterior sample of all of the chains. Another suggestion is that we could split the entire area into a grid and reduce the number of locations to the number of centers of the grid squares. However, this approach would probably reduce the spatial effect of the data and based on the location distribution of some moose we could have a lot of empty squares with no observation.

Another suggestion is that we could specify the spatial effect for the individual covariates. It would be possible to set up more candidate models based on which of the covariates vary spatially and perform more comparisons to find out which covariates bring us the significant difference and for which it is unnecessary to include a spatial effect. Also we could then obtain the best model for each moose individually to receive the best model fit performance.

Finally, we could obtain covariates for other locations in the study area or leave some portion of the data as a validation sample and assess the model predictive performance.

References

[1] Banerjee, Sudipto, Bradley P. Carlin, and Alan E. Gelfand. Hierarchical Modeling and Analysis for Spatial Data. Second ed. Boca Raton: CRC, Taylor and Francis Group, 2015. Print.

[2] Hobbs, N. Thompson., and Mevin B. Hooten. Bayesian Models: A Statistical Primer for Ecologists. Princeton, NJ: Princeton UP, 2015. Print.

[3] Banerjee, Sudipto, Alan E. Gelfand, Andrew O. Finley, and Huiyan Sang. "Gaussian Predictive Process Models for Large Spatial Data Sets." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 70.4 (2008): 825-48. Web.

[4] Babcock, Chad, Andrew O. Finley, John B. Bradford, Randall Kolka, Richard Birdsey, and Michael G. Ryan. "LiDAR Based Prediction of Forest Biomass Using Hierarchical Models with Spatially Varying Coefficients." Remote Sensing of Environment 169 (2015): 113-27. Web.

[5] Babcock, Chad, Jason Matney, Andrew O. Finley, Aaron Weiskittel, and Bruce D. Cook. "Multivariate Spatial Regression Models for Predicting Individual Tree Structure Variables Using LiDAR Data." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6.1 (2013): 6-14. Web.

[6] Lawson, Andrew, Sudipto Banerjee, Robert P. Haining, and Maria Dolores Ugarte. Handbook of Spatial Epidemiology. Boca Raton: CRC/Taylor and Francis, 2016. Print.

[7] Bakar, K. Shuvo, Philip Kokic, and Huidong Jin. "Hierarchical Spatially Varying Coefficient and Temporal Dynamic Process Models UsingspTDyn." Journal of Statistical Computation and Simulation 86.4 (2015): 820-40. Web.

[8] Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park. "Package 'MCMCpack'." (2017): 79-83 Web.

[9] Finley, Andrew O., and Sudipto Banerjee. "Package 'spBayes'." (2017): 32-37 Web.

[10] Finley, Andrew O., Sudipto Banerjee, and Alan E.gelfand. "SpBayesfor Large Univariate and Multivariate Point-Referenced Spatio-Temporal Data Models." Journal of Statistical Software 63.13 (2015): 1-28. Web.

[11] Roberts, Gareth O., and Jeffrey S. Rosenthal. "Coupling and Ergodicity of Adaptive Markov Chain Monte Carlo Algorithms." Journal of Applied Probability 44.02 (2007): 458-75. Web.

[12] Banerjee, Sudipto, Bradley P. Carlin, and Alan E. Gelfand. "5.2 Generalized Linearspatial Process Modeling." Hierarchical Modeling and Analysis for Spatial Data. First ed. Boca Raton: CRC, Taylor and Francis Group, 2004. 159-162 Print.

[13] McCulloch, Charles E., and Shayle R. Searle. "Generalized Linear Models (GLMs)." Generalized, Linear, and Mixed Models Wiley Series in Probability and Statistics (2005): 135-55. Web.

[14] Gelman, Andrew, Jessica Hwang, and Aki Vehtari. "Understanding Predictive Information Criteria for Bayesian Models." Statistics and Computing 24.6 (2013): 997-1016. Web.

[15] "Moose." National Geographic. National Geographic, 24 May 2017. Web. 10 June 2017. http://www.nationalgeographic.com/animals/mammals/m/moose/.

[16] DelGiudice, Glenn D. "2017 Aerial Moose Survey." Moose Population Survey (n.d.): n. pag. Minnesota Department of Natural Resources. Web. http://files.dnr.state.mn.us/wildlife/moose/moosesurvey.pdf.

[17] Robbins, Jim. "Moose Die-Off Alarms Scientists." The New York Times. The New York Times, 14 Oct. 2013. Web. 10 June 2017. http://www.nytimes.com/2013/10/15/science/earth/something-is-killing-off-the-moose.html?_r=0.

Appendices

Appendix I - Tables for the non-spatial models using all observations

Moose ID 31168

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.964	0.317	-9.338	<0.001	0.052	0.028	0.096
Collar Temperature	-0.024	0.013	-1.820	0.069	0.976	0.952	1.002
Open Water	5.461	0.191	28.614	<0.001	235.359	161.908	342.129
Mixedwood Forest	-0.436	0.184	-2.363	0.018	0.647	0.451	0.928
Low Density Forest	-15.467	278.441	-0.056	0.956	0.000	0.000	N/A
Time Day/Night	-0.915	0.160	-5.710	<0.001	0.401	0.293	0.548
04/15 - 05/31	-0.536	0.212	-2.530	0.011	0.585	0.386	0.886
08/01 - 10/31	-0.773	0.193	-4.012	<0.001	0.462	0.317	0.674

Table 16: Generalized linear model, Moose ID: 31168

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.983	0.316	-9.443	<0.001	0.051	0.027	0.094
Collar Temperature	-0.023	0.013	-1.828	0.068	0.977	0.953	1.002
Open Water	5.495	0.196	28.105	<0.001	243.374	165.907	357.013
Mixedwood Forest	-0.439	0.188	-2.332	0.020	0.644	0.446	0.932
Low Density Forest	-326.167	204.967	-1.591	0.112	0.000	0.000	N/A
Time Day/Night	-0.920	0.170	-5.407	<0.001	0.399	0.286	0.556
04/15 - 05/31	-0.531	0.192	-2.761	0.006	0.588	0.403	0.857
08/01 - 10/31	-0.775	0.182	-4.267	<0.001	0.461	0.323	0.658

Table 17: Bayesian Non-spatial model (MCMCpack), Moose ID: 31168

Moose ID 31169

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.034	0.181	-16.739	<0.001	0.048	0.034	0.069
Collar Temperature	0.044	0.008	5.294	<0.001	1.045	1.028	1.063
Open Water	1.383	0.209	6.618	<0.001	3.988	2.647	6.007
Mixedwood Forest	-2.099	0.115	-18.190	<0.001	0.123	0.098	0.154
Low Density Forest	-3.397	0.296	-11.472	<0.001	0.033	0.019	0.060
Time Day/Night	-0.779	0.120	-6.500	<0.001	0.459	0.363	0.580
04/15 - 05/31	-0.625	0.167	-3.737	<0.001	0.535	0.386	0.743
08/01 - 10/31	-0.269	0.116	-2.313	0.021	0.764	0.608	0.960

Table 18: Generalized linear model, Moose ID: 31169

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.035	0.182	-16.676	<0.001	0.048	0.034	0.069
Collar Temperature	0.044	0.008	5.216	<0.001	1.045	1.028	1.062
Open Water	1.394	0.219	6.360	<0.001	4.032	2.624	6.195
Mixedwood Forest	-2.100	0.119	-17.650	<0.001	0.122	0.097	0.155
Low Density Forest	-3.434	0.305	-11.246	<0.001	0.032	0.018	0.059
Time Day/Night	-0.771	0.118	-6.524	<0.001	0.463	0.367	0.583
04/15 - 05/31	-0.622	0.158	-3.931	<0.001	0.537	0.394	0.732
08/01 - 10/31	-0.271	0.116	-2.330	0.020	0.763	0.607	0.958

Table 19: Bayesian Non-spatial model (MCMCpack), Moose ID: 31169

Moose ID 31170

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.677	0.566	-4.728	<0.001	0.069	0.023	0.209
Collar Temperature	0.029	0.007	4.197	<0.001	1.030	1.016	1.044
Open Water	2.005	0.539	3.721	<0.001	7.425	2.583	21.343
Mixedwood Forest	-1.775	0.554	-3.204	0.001	0.169	0.057	0.502
Low Density Forest	-1.867	0.541	-3.451	0.001	0.155	0.054	0.446
Time Day/Night	0.395	0.100	3.948	<0.001	1.484	1.220	1.806
04/15 - 05/31	-1.447	0.228	-6.356	<0.001	0.235	0.150	0.367
08/01 - 10/31	-0.057	0.121	-0.467	0.640	0.945	0.745	1.198

Table 20: Generalized linear model, Moose ID: 31170

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.746	0.591	-4.644	<0.001	0.064	0.020	0.205
Collar Temperature	0.029	0.007	4.047	<0.001	1.029	1.015	1.044
Open Water	2.077	0.547	3.796	<0.001	7.981	2.731	23.320
Mixedwood Forest	-1.709	0.578	-2.954	0.003	0.181	0.058	0.563
Low Density Forest	-1.789	0.559	-3.201	0.001	0.167	0.056	0.500
Time Day/Night	0.394	0.098	3.998	<0.001	1.483	1.222	1.798
04/15 - 05/31	-1.445	0.218	-6.633	<0.001	0.236	0.154	0.361
08/01 - 10/31	-0.058	0.118	-0.492	0.622	0.943	0.748	1.190

Table 21: Bayesian Non-spatial model (MCMCpack), Moose ID: 31170

Moose ID 31172

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.003	0.277	-10.837	<0.001	0.050	0.029	0.085
Collar Temperature	0.052	0.013	4.032	<0.001	1.053	1.027	1.080
Open Water	-13.249	387.969	-0.034	0.973	0.000	0.000	N/A
Mixedwood Forest	-3.026	0.188	-16.106	<0.001	0.048	0.034	0.070
Low Density Forest	-1.624	0.194	-8.356	<0.001	0.197	0.135	0.289
Time Day/Night	-0.020	0.186	-0.105	0.916	0.981	0.681	1.412
04/15 - 05/31	-0.271	0.209	-1.299	0.194	0.763	0.507	1.148
08/01 - 10/31	-0.100	0.197	-0.509	0.611	0.905	0.615	1.331

Table 22: Generalized linear model, Moose ID: 31172

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.037	0.277	-10.984	<0.001	0.048	0.028	0.082
Collar Temperature	0.053	0.013	3.986	<0.001	1.054	1.027	1.082
Open Water	-356.259	202.719	-1.757	0.079	0.000	0.000	N/A
Mixedwood Forest	-3.033	0.198	-15.286	<0.001	0.048	0.033	0.071
Low Density Forest	-1.622	0.201	-8.069	<0.001	0.198	0.133	0.293
Time Day/Night	-0.033	0.188	-0.175	0.861	0.968	0.670	1.398
04/15 - 05/31	-0.269	0.196	-1.376	0.169	0.764	0.521	1.121
08/01 - 10/31	-0.106	0.186	-0.568	0.570	0.900	0.624	1.296

Table 23: Bayesian Non-spatial model (MCMCpack), Moose ID: 31172

Moose ID 31174

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.569	0.322	-11.072	<0.001	0.028	0.015	0.053
Collar Temperature	0.041	0.015	2.788	0.005	1.042	1.012	1.072
Open Water	5.151	0.252	20.461	<0.001	172.564	105.358	282.639
Mixedwood Forest	-2.840	0.221	-12.830	<0.001	0.058	0.038	0.090
Low Density Forest	-4.722	1.006	-4.692	<0.001	0.009	0.001	0.064
Time Day/Night	-0.890	0.189	-4.716	<0.001	0.411	0.284	0.594
04/15 - 05/31	-1.449	0.408	-3.550	<0.001	0.235	0.106	0.523
08/01 - 10/31	-0.571	0.186	-3.066	0.002	0.565	0.392	0.814

Table 24: Generalized linear model, Moose ID: 31174

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.613	0.336	-10.762	<0.001	0.027	0.014	0.052
Collar Temperature	0.042	0.015	2.742	0.006	1.043	1.012	1.074
Open Water	5.191	0.272	19.107	<0.001	179.630	105.471	305.932
Mixedwood Forest	-2.842	0.227	-12.513	<0.001	0.058	0.037	0.091
Low Density Forest	-5.322	1.491	-3.569	<0.001	0.005	0.000	0.091
Time Day/Night	-0.886	0.187	-4.738	<0.001	0.412	0.286	0.595
04/15 - 05/31	-1.479	0.411	-3.595	<0.001	0.228	0.102	0.510
08/01 - 10/31	-0.564	0.186	-3.036	0.002	0.569	0.396	0.819

Table 25: Bayesian Non-spatial model (MCMCpack), Moose ID: 31174

Moose ID 31178

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.337	0.484	-8.959	<0.001	0.013	0.005	0.034
Collar Temperature	-0.034	0.019	-1.801	0.072	0.966	0.931	1.003
Open Water	5.773	0.486	11.881	<0.001	321.475	124.039	833.178
Mixedwood Forest	0.174	0.274	0.635	0.525	1.190	0.695	2.038
Low Density Forest	-2.720	0.743	-3.659	<0.001	0.066	0.015	0.283
Time Day/Night	-0.946	0.247	-3.834	<0.001	0.388	0.240	0.630
04/15 - 05/31	-1.910	0.478	-3.999	<0.001	0.148	0.058	0.378
08/01 - 10/31	-0.183	0.275	-0.668	0.504	0.832	0.486	1.426

Table 26: Generalized linear model, Moose ID: 31178

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.302	0.508	-8.468	<0.001	0.014	0.005	0.037
Collar Temperature	-0.037	0.020	-1.849	0.065	0.963	0.926	1.002
Open Water	5.838	0.469	12.440	<0.001	343.068	136.745	860.696
Mixedwood Forest	0.169	0.269	0.628	0.530	1.184	0.699	2.005
Low Density Forest	-2.966	0.891	-3.328	0.001	0.052	0.009	0.296
Time Day/Night	-0.943	0.236	-3.989	<0.001	0.389	0.245	0.619
04/15 - 05/31	-1.993	0.454	-4.386	<0.001	0.136	0.056	0.332
08/01 - 10/31	-0.214	0.283	-0.757	0.449	0.807	0.463	1.406

Table 27: Bayesian Non-spatial model (MCMCpack), Moose ID: 31178

Moose ID 31182

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-6.316	1.128	-5.602	<0.001	0.002	0.000	0.016
Collar Temperature	0.037	0.044	0.851	0.395	1.038	0.952	1.132
Open Water	5.389	1.090	4.944	<0.001	218.938	25.851	1854.238
Mixedwood Forest	-17.563	965.339	-0.018	0.986	0.000	0.000	N/A
Low Density Forest	-17.488	1694.905	-0.010	0.992	0.000	0.000	N/A
Time Day/Night	-1.457	0.670	-2.176	0.030	0.233	0.063	0.866
04/15 - 05/31	0.197	0.990	0.199	0.842	1.218	0.175	8.469
08/01 - 10/31	0.541	0.814	0.665	0.506	1.718	0.348	8.468

Table 28: Generalized linear model, Moose ID: 31182

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-7.134	0.935	-7.630	<0.001	0.001	0.000	0.005
Collar Temperature	0.052	0.042	1.244	0.214	1.053	0.970	1.143
Open Water	6.165	1.250	4.933	<0.001	476.020	41.102	5513.011
Mixedwood Forest	-374.853	195.637	-1.916	0.055	0.000	0.000	5365.182
Low Density Forest	-312.230	177.134	-1.763	0.078	0.000	0.000	N/A
Time Day/Night	-1.370	0.788	-1.739	0.082	0.254	0.054	1.190
04/15 - 05/31	0.669	0.919	0.728	0.466	1.953	0.322	11.830
08/01 - 10/31	0.886	0.846	1.046	0.295	2.424	0.461	12.738

Table 29: Bayesian Non-spatial model (MCMCpack), Moose ID: 31182

Moose ID 31184

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.554	0.644	-7.076	<0.001	0.011	0.003	0.037
Collar Temperature	0.032	0.012	2.658	0.008	1.032	1.008	1.057
Open Water	5.261	0.606	8.681	<0.001	192.757	58.765	632.263
Mixedwood Forest	-0.935	0.589	-1.587	0.112	0.393	0.124	1.245
Low Density Forest	0.570	0.623	0.914	0.360	1.768	0.521	5.994
Time Day/Night	-0.150	0.165	-0.904	0.366	0.861	0.623	1.191
04/15 - 05/31	-2.180	0.395	-5.523	<0.001	0.113	0.052	0.245
08/01 - 10/31	-0.395	0.169	-2.333	0.020	0.674	0.483	0.939

Table 30: Generalized linear model, Moose ID: 31184

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.791	0.676	-7.086	<0.001	0.008	0.002	0.031
Collar Temperature	0.032	0.012	2.617	0.009	1.032	1.008	1.058
Open Water	5.509	0.626	8.794	<0.001	246.815	72.308	842.473
Mixedwood Forest	-0.712	0.615	-1.158	0.247	0.491	0.147	1.637
Low Density Forest	0.803	0.644	1.248	0.212	2.232	0.632	7.881
Time Day/Night	-0.155	0.161	-0.964	0.335	0.856	0.625	1.174
04/15 - 05/31	-2.228	0.390	-5.716	<0.001	0.108	0.050	0.231
08/01 - 10/31	-0.408	0.165	-2.470	0.014	0.665	0.481	0.919

Table 31: Bayesian Non-spatial model (MCMCpack), Moose ID: 31184

Appendix II - Tables for the spatial model using selected observations

Moose ID 31168

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.118	0.549	-5.674	<0.001	0.044	0.015	0.130
Collar Temperature	-0.004	0.022	-0.162	0.871	0.996	0.954	1.041
Open Water	5.582	0.328	17.005	<0.001	265.675	139.612	505.564
Mixedwood Forest	-0.755	0.309	-2.448	0.014	0.470	0.257	0.860
Low Density Forest	-15.678	470.733	-0.033	0.973	0.000	0.000	N/A
Time Day/Night	-0.986	0.276	-3.572	<0.001	0.373	0.217	0.641
04/15 - 05/31	-0.299	0.367	-0.814	0.416	0.742	0.362	1.522
08/01 - 10/31	-0.801	0.336	-2.385	0.017	0.449	0.233	0.867

Table 32: Generalize linear model for moose ID 31168 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-5.308	1.163	-4.564	<0.001	0.005	0.001	0.048
Collar Temperature	0.019	0.037	0.509	0.610	1.019	0.948	1.096
Open Water	5.477	0.990	5.532	<0.001	239.128	34.351	1664.640
Mixedwood Forest	-3.287	0.821	-4.002	<0.001	0.037	0.007	0.187
Low Density Forest	-349.400	197.800	-1.766	0.077	0.000	0.000	N/A
Time Day/Night	-1.162	0.464	-2.506	0.012	0.313	0.126	0.776
04/15 - 05/31	-1.978	0.902	-2.193	0.028	0.138	0.024	0.811
08/01 - 10/31	-1.645	0.692	-2.376	0.017	0.193	0.050	0.750
sigma.sq	25.640	8.794	2.916	0.004			
phi	0.003	0.001	3.799	<0.001			

Table 33: Bayesian model with spatial effect for moose ID 31168 with $n = 5,000$.

Moose ID 31169

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.147	0.447	-7.041	<0.001	0.043	0.018	0.103
Collar Temperature	0.066	0.021	3.092	0.002	1.068	1.024	1.114
Open Water	1.470	0.439	3.346	0.001	4.349	1.838	10.289
Mixedwood Forest	-2.251	0.286	-7.867	<0.001	0.105	0.060	0.184
Low Density Forest	-3.150	0.598	-5.264	<0.001	0.043	0.013	0.138
Time Day/Night	-1.196	0.295	-4.059	<0.001	0.302	0.170	0.539
04/15 - 05/31	-0.877	0.436	-2.011	0.044	0.416	0.177	0.978
08/01 - 10/31	-0.334	0.273	-1.223	0.221	0.716	0.420	1.223

Table 34: Generalize linear model for moose ID 31169 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-9.042	1.556	-5.809	<0.001	0.000	0.000	0.003
Collar Temperature	0.051	0.031	1.635	0.102	1.052	0.990	1.118
Open Water	0.914	1.316	0.694	0.487	2.494	0.189	32.899
Mixedwood Forest	-1.233	0.716	-1.723	0.085	0.291	0.072	1.185
Low Density Forest	-1.764	1.568	-1.125	0.261	0.171	0.008	3.705
Time Day/Night	-0.794	0.469	-1.692	0.091	0.452	0.180	1.134
04/15 - 05/31	-2.091	0.929	-2.251	0.024	0.124	0.020	0.763
08/01 - 10/31	0.058	0.586	0.099	0.921	1.060	0.336	3.345
sigma.sq	32.218	13.067	2.466	0.014			
phi	0.007	0.002	3.403	0.001			

Table 35: Bayesian model with spatial effect for moose ID 31169 with $n = 5,000$.

Moose ID 31170

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-2.946	1.103	-2.671	0.008	0.053	0.006	0.457
Collar Temperature	0.028	0.012	2.303	0.021	1.028	1.004	1.053
Open Water	2.298	1.058	2.172	0.030	9.955	1.252	79.185
Mixedwood Forest	-1.599	1.081	-1.479	0.139	0.202	0.024	1.682
Low Density Forest	-1.708	1.061	-1.609	0.108	0.181	0.023	1.451
Time Day/Night	0.545	0.180	3.033	0.002	1.725	1.213	2.454
04/15 - 05/31	-1.331	0.369	-3.611	<0.001	0.264	0.128	0.544
08/01 - 10/31	-0.254	0.216	-1.176	0.240	0.776	0.509	1.184

Table 36: Generalize linear model for moose ID 31170 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-12.817	3.101	-4.133	<0.001	0.000	0.000	0.001
Collar Temperature	0.017	0.040	0.427	0.670	1.017	0.941	1.100
Open Water	2.706	1.860	1.455	0.146	14.973	0.391	573.371
Mixedwood Forest	-6.136	3.949	-1.554	0.120	0.002	0.000	4.973
Low Density Forest	-2.481	2.067	-1.201	0.230	0.084	0.001	4.803
Time Day/Night	0.884	0.506	1.749	0.080	2.422	0.899	6.524
04/15 - 05/31	-0.950	1.333	-0.713	0.476	0.387	0.028	5.274
08/01 - 10/31	1.311	0.956	1.372	0.170	3.711	0.570	24.150
sigma.sq	429.097	152.727	2.810	0.005			
phi	0.004	0.001	3.078	0.002			

Table 37: Bayesian model with spatial effect for moose ID 31170 with $n = 5,000$.

Moose ID 31172

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.444	0.504	-6.837	<0.001	0.032	0.012	0.086
Collar Temperature	0.053	0.022	2.399	0.016	1.054	1.010	1.101
Open Water	-12.737	548.207	-0.023	0.981	0.000	0.000	N/A
Mixedwood Forest	-2.754	0.339	-8.136	<0.001	0.064	0.033	0.124
Low Density Forest	-1.101	0.332	-3.314	0.001	0.333	0.173	0.638
Time Day/Night	0.244	0.339	0.721	0.471	1.277	0.657	2.479
04/15 - 05/31	-0.542	0.381	-1.423	0.155	0.582	0.276	1.227
08/01 - 10/31	-0.424	0.346	-1.226	0.220	0.654	0.332	1.289

Table 38: Generalize linear model for moose ID 31172 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-9.981	1.195	-8.352	<0.001	0.000	0.000	0.000
Collar Temperature	0.091	0.029	3.153	0.002	1.096	1.035	1.160
Open Water	-336.800	201.500	-1.671	0.095	0.000	0.000	N/A
Mixedwood Forest	-1.451	0.692	-2.098	0.036	0.234	0.060	0.909
Low Density Forest	-1.431	0.893	-1.603	0.109	0.239	0.042	1.376
Time Day/Night	0.325	0.436	0.746	0.456	1.384	0.589	3.254
04/15 - 05/31	-0.613	0.827	-0.741	0.459	0.542	0.107	2.740
08/01 - 10/31	0.760	0.611	1.243	0.214	2.137	0.645	7.081
sigma.sq	23.720	6.428	3.690	<0.001			
phi	0.004	0.001	3.479	0.001			

Table 39: Bayesian model with spatial effect for moose ID 31172 with $n = 5,000$.

Moose ID 31174

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.075	0.797	-5.110	<0.001	0.017	0.004	0.081
Collar Temperature	0.066	0.037	1.792	0.073	1.068	0.994	1.148
Open Water	4.665	0.549	8.500	<0.001	106.196	36.218	311.387
Mixedwood Forest	-2.888	0.556	-5.193	<0.001	0.056	0.019	0.166
Low Density Forest	-17.200	733.543	-0.023	0.981	0.000	0.000	N/A
Time Day/Night	-1.156	0.477	-2.425	0.015	0.315	0.124	0.801
04/15 - 05/31	-1.622	1.086	-1.494	0.135	0.197	0.024	1.659
08/01 - 10/31	-0.358	0.437	-0.819	0.413	0.699	0.297	1.646

Table 40: Generalize linear model for moose ID 31174 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-8.462	1.587	-5.332	<0.001	0.000	0.000	0.005
Collar Temperature	0.111	0.049	2.244	0.025	1.117	1.014	1.231
Open Water	7.019	1.407	4.989	<0.001	1117.668	70.904	17617.876
Mixedwood Forest	-2.049	0.812	-2.523	0.012	0.129	0.026	0.633
Low Density Forest	-346.900	195.600	-1.774	0.076	0.000	0.000	N/A
Time Day/Night	-0.497	0.620	-0.802	0.423	0.608	0.180	2.051
04/15 - 05/31	-1.714	1.479	-1.159	0.247	0.180	0.010	3.270
08/01 - 10/31	-0.063	0.714	-0.088	0.930	0.939	0.232	3.805
sigma.sq	10.250	5.142	1.993	0.046			
phi	0.007	0.005	1.452	0.147			

Table 41: Bayesian model with spatial effect for moose ID 31174 with $n = 5,000$.

Moose ID 31178

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.193	1.060	-3.957	<0.001	0.015	0.002	0.120
Collar Temperature	0.008	0.044	0.192	0.848	1.008	0.925	1.099
Open Water	5.322	1.258	4.231	<0.001	204.741	17.398	2409.461
Mixedwood Forest	-0.576	0.565	-1.020	0.308	0.562	0.186	1.700
Low Density Forest	-16.680	824.389	-0.020	0.984	0.000	0.000	N/A
Time Day/Night	-1.306	0.590	-2.213	0.027	0.271	0.085	0.861
04/15 - 05/31	-1.914	1.135	-1.687	0.092	0.147	0.016	1.364
08/01 - 10/31	-0.294	0.614	-0.479	0.632	0.745	0.224	2.482

Table 42: Generalize linear model for moose ID 31178 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-5.573	2.000	-2.786	0.005	0.004	0.000	0.192
Collar Temperature	0.015	0.052	0.296	0.768	1.015	0.917	1.124
Open Water	6.875	2.515	2.733	0.006	967.475	6.992	133862.417
Mixedwood Forest	-0.500	0.709	-0.706	0.480	0.606	0.151	2.433
Low Density Forest	-348.072	198.707	-1.752	0.080	0.000	0.000	N/A
Time Day/Night	-1.365	0.664	-2.055	0.040	0.255	0.069	0.939
04/15 - 05/31	-2.042	1.466	-1.394	0.163	0.130	0.007	2.294
08/01 - 10/31	0.063	0.861	0.073	0.942	1.065	0.197	5.761
sigma.sq	7.130	8.151	0.875	0.382			
phi	0.716	0.952	0.751	0.453			

Table 43: Bayesian model with spatial effect for moose ID 31178 with $n = 5,000$.

Moose ID 31179

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	0.327	0.588	0.556	0.578	1.386	0.438	4.390
Collar Temperature	0.026	0.013	2.028	0.043	1.026	1.001	1.053
Open Water	-0.868	0.543	-1.598	0.110	0.420	0.145	1.217
Mixedwood Forest	-3.626	0.553	-6.555	<0.001	0.027	0.009	0.079
Low Density Forest	-4.546	0.553	-8.221	<0.001	0.011	0.004	0.031
Time Day/Night	0.022	0.161	0.139	0.889	1.023	0.746	1.402
04/15 - 05/31	-0.923	0.244	-3.780	<0.001	0.397	0.246	0.641
08/01 - 10/31	-0.913	0.180	-5.085	<0.001	0.401	0.282	0.571

Table 44: Generalize linear model for moose ID 31179 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-23.407	3.001	-7.800	<0.001	0.000	0.000	0.000
Collar Temperature	0.086	0.034	2.519	0.012	1.090	1.019	1.166
Open Water	2.273	3.335	0.682	0.495	9.712	0.014	6699.686
Mixedwood Forest	-4.883	4.092	-1.193	0.233	0.008	0.000	23.040
Low Density Forest	-1.826	3.327	-0.549	0.583	0.161	0.000	109.340
Time Day/Night	-0.023	0.386	-0.059	0.953	0.978	0.459	2.083
04/15 - 05/31	1.522	1.740	0.875	0.382	4.582	0.151	138.729
08/01 - 10/31	0.584	0.625	0.933	0.351	1.793	0.526	6.108
sigma.sq	501.816	215.600	2.328	0.020			
phi	0.003	0.001	3.859	<0.001			

Table 45: Bayesian model with spatial effect for moose ID 31179 with $n = 5,000$.

Moose ID 31182

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-8.222	2.646	-3.107	0.002	0.000	0.000	0.048
Collar Temperature	0.200	0.118	1.699	0.089	1.221	0.970	1.538
Open Water	5.624	2.063	2.726	0.006	276.967	4.858	15789.646
Mixedwood Forest	-18.880	3884.999	-0.005	0.996	0.000	0.000	N/A
Low Density Forest	-19.794	7307.065	-0.003	0.998	0.000	0.000	N/A
Time Day/Night	-3.566	1.824	-1.955	0.051	0.028	0.001	1.010
04/15 - 05/31	-17.905	5251.840	-0.003	0.997	0.000	0.000	N/A
08/01 - 10/31	-1.187	1.319	-0.900	0.368	0.305	0.023	4.049

Table 46: Generalize linear model for moose ID 31182 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-8.841	2.570	-3.440	0.001	0.000	0.000	0.022
Collar Temperature	0.205	0.114	1.793	0.073	1.227	0.981	1.535
Open Water	6.175	2.330	2.650	0.008	480.679	4.996	46245.530
Mixedwood Forest	-261.502	169.280	-1.545	0.122	0.000	0.000	N/A
Low Density Forest	-265.338	171.413	-1.548	0.122	0.000	0.000	N/A
Time Day/Night	-4.001	2.095	-1.910	0.056	0.018	0.000	1.111
04/15 - 05/31	-170.824	133.351	-1.281	0.200	0.000	0.000	N/A
08/01 - 10/31	-1.405	1.592	-0.882	0.378	0.245	0.011	5.563
sigma.sq	2.073	1.887	1.099	0.272			
phi	1.414	0.903	1.565	0.118			

Table 47: Bayesian model with spatial effect for moose ID 31182 with $n = 5,000$.

Moose ID 31184

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-3.727	0.987	-3.776	<0.001	0.024	0.003	0.167
Collar Temperature	0.065	0.028	2.299	0.021	1.067	1.010	1.127
Open Water	4.574	0.872	5.246	<0.001	96.947	17.551	535.505
Mixedwood Forest	-2.017	0.760	-2.653	0.008	0.133	0.030	0.590
Low Density Forest	-0.665	0.893	-0.745	0.456	0.514	0.089	2.957
Time Day/Night	-0.612	0.393	-1.555	0.120	0.542	0.251	1.173
04/15 - 05/31	-2.600	1.094	-2.377	0.017	0.074	0.009	0.634
08/01 - 10/31	-0.552	0.387	-1.427	0.154	0.576	0.270	1.229

Table 48: Generalize linear model for moose ID 31184 with $n = 5,000$.

	Estimate	Std. Error	z value	Pr(>|z|)	Odds ratio	L 0.95 CI	U 0.95 CI
Intercept	-4.344	1.181	-3.678	<0.001	0.013	0.001	0.131
Collar Temperature	0.053	0.032	1.654	0.098	1.054	0.990	1.123
Open Water	4.387	1.236	3.550	<0.001	80.378	7.133	905.706
Mixedwood Forest	-2.962	0.999	-2.964	0.003	0.052	0.007	0.367
Low Density Forest	-1.755	1.203	-1.459	0.144	0.173	0.016	1.826
Time Day/Night	-0.607	0.455	-1.334	0.182	0.545	0.223	1.330
04/15 - 05/31	-3.861	1.589	-2.430	0.015	0.021	0.001	0.474
08/01 - 10/31	-0.616	0.478	-1.289	0.197	0.540	0.212	1.378
sigma.sq	7.828	2.834	2.762	0.006			
phi	0.008	0.003	3.059	0.002			

Table 49: Bayesian model with spatial effect for moose ID 31184 with $n = 5,000$.

Appendix III - R code

```
#load data
setwd("C:/users/mathgrad/Desktop/Matej/R")
Data <- read.csv(file="C:/users/mathgrad/Desktop/Matej/data/
    AlcesPaperData_clear_new.csv", header=TRUE, sep=",")

#preparing the data set
Data[is.na(Data)] <- 0
Data <- Data[complete.cases(Data), ]

#elimination of low-frequency class code observations
el <- which(Data$CLASS_CODE==10 | Data$CLASS_CODE==11 | Data$CLASS_CODE==12)
if (length(el) != 0){
  Data <- Data[-el, ]
}

#make a sample
#select moose by ID
Data <- subset(Data, Moose_ID==31168) #available ID's: 31166, 31168, 31169,
    31172, 31174, 31178, 31179, 31182, 31184

#Use all datapoints
Data_fit <- Data
n <- nrow(Data_fit)

#random sample of size n
n <- 5000
sample <- sample(nrow(Data), n)
Data_fit <- Data[sample, ]

#first n observations
n <- 8000
Data_fit <- Data[(1:n), ]

#creating matrix of coordinates
coords <- cbind(Data_fit[, "UTMx"], Data_fit[, "UTMy"])
colnames(coords) <- c("UTMx", "UTMy")

```

```
#computing knots based on K-means and plot
par(mfrow=c(1,1))
m <- floor(0.1*n)
km.knots <- kmeans(coords, m)$centers
plot(coords, pch=19, cex=0.5, xlab="Easting (m)", ylab="Northing (m)")
points(km.knots, pch=19, cex=0.5, col="red")

#factor and references
Data_fit$CLASS_CODE_NEW <- factor(Data_fit$CLASS_CODE_NEW)
Data_fit <- within(Data_fit, CLASS_CODE_NEW <- relevel(CLASS_CODE_NEW, ref = 2))

Data_fit$Date_NEW <- factor(Data_fit$Date_NEW)
Data_fit <- within(Data_fit, Date_NEW <- relevel(Date_NEW, ref = 2))

#GLM Analysis - choosing the covariates
formula <- as.formula(Hydro_pom ~ Collar_Temp + CLASS_CODE_NEW + Time_NEW + Date_NEW)
# #backward selection
# fit <- glm(formula, family="binomial", data=Data_fit)
# AIC <- step(fit, direction = "backward", trace = 1)
# if (length(AIC$coefficients)!=8){
#     formula <- AIC$formula
# }

#Collect samples for Beta
fit <- glm(formula, family="binomial", data=Data_fit)
parameters <- colnames(fit$R)
summary(fit)
beta.starting <- coefficients(fit)

#binom response for GLM
weights <- rep(1, n) #the number of trials in each location
p.hat <- as.matrix(fit$fitted.values)
glm.binom <- apply(p.hat, 2, function(x){rbinom(n, size=weights, prob=p.hat)})



```

```r
#transformation of glm fit
y.hat.glm.new <- matrix(rep(0,n))
norm <- 1/max(fit$fitted.values)
for (i in 1:n){
  y.hat.glm.new[i] <- fit$fitted.values[i]*norm
}

#MSE for GLM
diff <- data.frame()
for (i in 1:n){
  diff[i,1] <- (Data_fit$Hydro_pom[i]-glm.binom[i])^2
}
mse_glm.binom <- sum(diff)/n

diff <- data.frame()
for (i in 1:n){
  diff[i,1] <- (Data_fit$Hydro_pom[i]-fit$fitted.values[i])^2
}
mse_glm <- sum(diff)/n

diff <- data.frame()
for (i in 1:n){
  diff[i,1] <- (Data_fit$Hydro_pom[i]-y.hat.glm.new[i])^2
}
mse_glm.new <- sum(diff)/n

print(mse_glm.binom)
print(mse_glm)
print(mse_glm.new)

##########################################################################
#Bayes non-spatial regression model by MCMCpack package
library(MCMCpack)
iterations <- 100000
burnin <- iterations*0.75
mcmc <- iterations-burnin

MCMCpack_model <- MCMClogit(formula, data = Data_fit, burnin = burnin, mcmc
    = mcmc, thin = 1, verbose = 1, beta.start = 0, mubeta = 0, Vbeta = Inf)
```

```
summary(MCMCpack_model)
###############################################################################
library(spBayes)

#Adaptive Metropolis parameters
n.batch <- 3000
batch.length <- 10
n.samples <- n.batch*batch.length
#burn-in
burn.in <- floor(0.75*n.samples)
sub.samps <- burn.in:n.samples

#Bayes non-spatial regresion model by spBayes
weights <- rep(1, n) #the number of trials in each location
starting <- list("beta"=c(rep(0,length(parameters))))#starting values
priors <- list("beta.Flat") #prior distribution

#MCMC
Hydro.nonspGLM <- spGLM(formula, family="binomial", weights=weights, data=
    Data_fit,starting=starting, priors=priors, cov.model="exponential",
    verbose=TRUE, n.report=10, amcmc=list("n.batch"=n.batch, "batch.length"=
    batch.length, "accept.rate"=0.30))

print(summary(window(Hydro.nonspGLM$p.beta.samples, start=burn.in)))

###############################################################################
#univariate Bayesian spatial regression model by spBayes
weights <- rep(1, n) #the number of trials in each location
starting <- list("beta"=beta.starting, "phi"=3/50, "sigma.sq"=1, "w"=0)
priors <- list("beta.Flat", "phi.Unif"=c(0.00045, 3), "sigma.sq.IG"=c(2, 2))

#MCMC
Hydro.spGLM <- spGLM(formula, family="binomial", weights=weights, coords=
    coords,knots=km.knots, data=Data_fit, starting=starting, priors=priors,
    cov.model="exponential", verbose=TRUE, n.report=10, amcmc=list("n.batch"=
    n.batch, "batch.length"=batch.length, "accept.rate"=0.30))

print(summary(window(Hydro.spGLM$p.beta.theta.samples, start=burn.in)))
```

```r
################################################################################
#run time
print(Hydro.spGLM$run.time)

#Model fit diagnostics
print(spDiag(Hydro.spGLM, start = burn.in, thin = 1))
#plot(Hydro.spGLM$p.beta.theta.samples[sub.samps, "sigma.sq"])

#obtaining the samples for univariate Bayesian spatial regression model
beta.hat <- Hydro.spGLM$p.beta.theta.samples[sub.samps, parameters]
w.hat <- Hydro.spGLM$p.w.samples[, sub.samps]
beta.hat.transpose <- t(beta.hat)
p.hat <- 1/(1+exp(-(Hydro.spGLM$X%*%beta.hat.transpose+w.hat)))
p.hat.mean <- apply(p.hat, 1, mean)
y.hat <- apply(p.hat, 2, function(x){rbinom(n, size=weights, prob=p.hat.mean)})
y.hat.mu <- apply(y.hat, 1, mean)
y.hat.var <- apply(y.hat, 1, var)

#y.hat.binom consisting from 0's and 1's
sum <- 0
y.hat.binom <- data.frame(rep(0,n))
l <- n.samples-burn.in+1
for (i in 1:n){
  sum <- 0
  sum <- sum(y.hat[i, ])
   if(sum > (l-1)/2){
     y.hat.binom[i,1] <- 1
  }
}

#observed vs fitted plots
resolution <- 100
library(MBA)
library(fields)
par(mfrow=c(1,2))
surf_org <- mba.surf(cbind(coords,Data_fit$Hydro_pom), no.X=resolution, no.Y=resolution, extend=TRUE)$xyz.est
image.plot(surf_org, main="Hydro observed values")
```

```
#points(coords)

surf_fit <- mba.surf(cbind(coords,y.hat.binom), no.X=resolution, no.Y=
    resolution, extend=TRUE)$xyz.est
image.plot(surf_fit, main="Hydro fitted values")
#points(coords)

##Contour plots of mean and variance
par(mfrow=c(1,2))
surf <- mba.surf(cbind(coords,y.hat.mu),no.X=100, no.Y=100, extend=TRUE)$xyz
    .est
image(surf, main="Interpolated mean of posterior rate\n(observed rate)")
contour(surf, add=TRUE)
#text(coords, label=paste("(",Data_fit$Hydro_pom,")",sep=""))
surf <- mba.surf(cbind(coords,y.hat.var),no.X=100, no.Y=100, extend=TRUE)$
    xyz.est
image(surf, main="Interpolated variance of posterior rate\n(observed #
      of trials)")
contour(surf, add=TRUE)
#text(coords, label=paste("(",weights,")",sep=""))

#transformation of mu
y.hat.mu.new <- matrix(rep(0,n))
norm <- 1/max(y.hat.mu)
for (i in 1:n){
  y.hat.mu.new[i] <- y.hat.mu[i]*norm
}

#MSE for Bayes
diff <- data.frame()
for (i in 1:n){
  diff[i,1] <- (Data_fit$Hydro_pom[i]-y.hat.binom[i,1])^2
}
mse_binom <- sum(diff)/n

diff <- data.frame()
for (i in 1:n){
  diff[i,1] <- (Data_fit$Hydro_pom[i]-y.hat.mu[i])^2
}
```

```
mse_mu <- sum(diff)/n

diff <- data.frame()
for (i in 1:n){
  diff[i,1] <- (Data_fit$Hydro_pom[i]-y.hat.mu.new[i])^2
}
mse_mu.new <- sum(diff)/n

print(mse_binom)
print(mse_mu)
print(mse_mu.new)
```